

AFRL-SN-RS-TR-2005-362
Final Technical Report
October 2005



HIGH LINEARITY WIDEBAND COHERENT AMPLITUDE MODULATION (AM) RADIO FREQUENCY (RF)-PHOTONIC LINKS

HRL Laboratories LLC

APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.

**AIR FORCE RESEARCH LABORATORY
SENSORS DIRECTORATE
ROME RESEARCH SITE
ROME, NEW YORK**

STINFO FINAL REPORT

This report has been reviewed by the Air Force Research Laboratory, Information Directorate, Public Affairs Office (IFOIPA) and is releasable to the National Technical Information Service (NTIS). At NTIS it will be releasable to the general public, including foreign nations.

AFRL-SN-RS-TR-2005-362 has been reviewed and is approved for publication

APPROVED: /s/

BRIAN F. MCKEON
Project Engineer

FOR THE DIRECTOR: /s/

RICHARD G. SHAUGHNESSY, CHIEF
Rome Operations Office
Sensors Directorate

REPORT DOCUMENTATION PAGE			Form Approved OMB No. 074-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing this collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503				
1. AGENCY USE ONLY (Leave blank)		2. REPORT DATE OCTOBER 2005	3. REPORT TYPE AND DATES COVERED Final Jul 03 – Jul 05	
4. TITLE AND SUBTITLE HIGH LINEARITY WIDEBAND COHERENT AMPLITUDE MODULATION (AM) RADIO FREQUENCY (RF)-PHOTONIC LINKS			5. FUNDING NUMBERS C - F30602-03-C-0120 PE - 62500F PR - 517D TA - SN WU - 03	
6. AUTHOR(S) Robert R. Hayes				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) HRL Laboratories LLC 301 Malibu Canyon Road Malibu California 90285-4797			8. PERFORMING ORGANIZATION REPORT NUMBER N/A	
9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES) Air Force Research Laboratory/SNDP 25 Electronic Parkway Rome New York 13441-4515			10. SPONSORING / MONITORING AGENCY REPORT NUMBER AFRL-SN-RS-TR-2005-362	
11. SUPPLEMENTARY NOTES AFRL Project Engineer: Brian F. McKeon/SNDP/(315) 330-7348/ Brian.McKeon@rl.af.mil				
12a. DISTRIBUTION / AVAILABILITY STATEMENT APPROVED FOR PUBLIC RELEASE; DISTRIBUTION UNLIMITED.				12b. DISTRIBUTION CODE
13. ABSTRACT (Maximum 200 Words) The objective of this project was to develop a fiber-optic RF link with high Spur Free Dynamic Range (SFDR). The approach taken was to utilize coherent AM modulation techniques which would, theoretically, provide a 6 dB advantage over Intensity Modulation/Direct Detection methods, and then insert a linearized electro-optic modulator in an attempt to achieve an SFDR of 130 dB Hz. In the coherent AM portion, locked homodyne detection and COTS components were used to achieve a measured SFDR of > 120 dB Hz at a 10 GHz operating frequency. The link, which used polarization maintaining components throughout and fusion splices instead of connectors, had none of the instabilities historically associated with homodyne systems. Efforts were made to raise the SFDR by using a Y-fed linearized directional coupler modulator; extensive measurements and theoretical analyses were made on these devices. The modulators showed the linearization predicted by theory; a logarithmic slope of 4.6 was demonstrated. Some shortcomings were also observed and analyzed that shed light on the difficulties of achieving wideband operation. The present devices can be designed to operate at specific center frequencies with maximum performance, but the linearized bandwidth is less than 2 GHz.				
14. SUBJECT TERMS Photonic RF Link, Linearized Modulator, Coherent AM, Amplitude Modulation, Electro-Optic Modulator, Directional Coupler Modulator, Wideband, Spur Free Dynamic Range, SFDR				15. NUMBER OF PAGES 82
				16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT UNCLASSIFIED	18. SECURITY CLASSIFICATION OF THIS PAGE UNCLASSIFIED	19. SECURITY CLASSIFICATION OF ABSTRACT UNCLASSIFIED	20. LIMITATION OF ABSTRACT UL	

TABLE OF CONTENTS

EXECUTIVE SUMMARY	1
1.0 Introduction	2
2.0 Coherent AM Links	2
2.1 Overview.....	2
2.2 Theoretical Performance of an Unlinearized AM Link.....	4
2.3 Link Gain and Noise Figure	7
2.4 Laser Noise.....	8
2.5 Finite Extinction	8
2.6 The Use of an Optical Amplifier to Increase the SFDR.....	9
3.0 SFDR Measurements for the Unlinearized AM link	11
3.1 Experimental Setup	11
3.2 Measurement Technique	15
3.3 Configurational Effects.....	16
3.4 SFDR Measurement Data	19
3.5 Photodiode Nonlinearity.....	23
4.0 Linearized AM Modulators.....	26
4.1 Introduction.....	26
4.2 The Synthesized Directional Coupler Modulator.....	27
4.2.1 Theory	27
4.2.2 SFDR	34
4.2.3 Manufacturing Tolerances.....	35
4.2.4 Phase Reversal Sections and Fabrication Approach	35
4.3 Dual-Y Directional Coupler Modulators.....	37
4.3.1 Introduction	37
4.3.2 Sensitivity Analysis.....	40
4.3.3 Velocity Mismatch	47
4.3.4 Electrode Attenuation.....	51
5.0 DEVICE MEASUREMENT	54
5.1 The PACT Program Modulators.....	54
5.2 Length Adjustment.....	55
5.3 Screening	58
5.4 Harmonic Measurements.....	59
5.5 The Fabry-Perot Modulator Effect.....	65
6.0 Summary and Conclusions	70
References.....	71
Appendix A: Laser Excess Noise (REN)	72
Appendix B: Effects of Finite Modulator Extinction	75

LIST OF FIGURES

FIGURE 1. AMPLITUDE VERSUS INTENSITY MODULATION.....	3
FIGURE 2. A SIMPLE AM LINK USING A MACH ZEHNDER MODULATOR AND HOMODYNE DETECTION.....	4
FIGURE 3. A DUAL-BALANCED DETECTOR.....	6
FIGURE 4. SNR VERSUS POSITION IN LINK FOR A SYSTEM USING AN OPTICAL AMPLIFIER OF GAIN G , FOLLOWED BY AN ATTENUATOR HAVING A VALUE OF A	10
FIGURE 5. A SIMPLE HOMODYNE AM LINK.....	11
FIGURE 6. THE EXPERIMENTAL LINK WITH DETAILS OF THE LOCKING TECHNIQUE SHOWN.....	12
FIGURE 7. EXPERIMENTAL SETUP AND SUPPORT EQUIPMENT.....	13
FIGURE 8. THE X-CUT (PUSH-PULL) MACH-ZEHNDER MODULATOR.....	13
FIGURE 9. PIEZOELECTRIC LINE STRETCHER WOUND WITH PM FIBER.....	14
FIGURE 10. THE DUAL-BALANCED PHOTODETECTOR.....	14
FIGURE 11. AN EARLIER SETUP SHOWING THE CONTROL ELECTRONICS (THE BLUE-YELLOW BOX).....	15
FIGURE 12. A SINGLE-SIDED Z-CUT MACH-ZEHNDER MODULATOR.....	17
FIGURE 13. A PUSH-PULL (X-CUT) MACH-ZEHNDER MODULATOR.....	18
FIGURE 14. THE FUNDAMENTAL (UPPER) AND IM3 (LOWER) TONES FOR A LASER POWER OF 100 mW AND A TOTAL LOCAL OSCILLATOR PHOTODETECTOR CURRENT OF 15 mA. THE RED LINE IS THE MEASURED NOISE IN A 1 HZ BANDWIDTH.....	20
FIGURE 15. MEASURED TONES AND NOISE FOR A LASER POWER OF 100 mW AND TOTAL LOCAL OSCILLATOR PHOTODETECTOR CURRENT OF 5 mA.....	21
FIGURE 16. LINK WITH AN OPTICAL AMPLIFIER (EDFA) TO BOOST THE LASER POWER.....	22
FIGURE 17. MEASURED TONES AND NOISE FOR THE LINK WITH 500 mW OF MOPA POWER AND A TOTAL LOCAL OSCILLATOR PHOTODETECTOR CURRENT OF 15 mA.....	22
FIGURE 18. TWO CLOSELY-SPACE OPTICAL WAVEGUIDES THAT FORM A DIRECTIONAL COUPLER.....	27
FIGURE 19. A TRIANGULAR MODULATOR TRANSFER CURVE.....	29
FIGURE 20. THE SHAPE OF $\kappa(z)$ NEEDED TO GET THE PROFILE OF FIGURE 19.....	30
FIGURE 21. THE TRANSFER CURVE THAT RESULTS FROM THE κ OF FIGURE 20.....	30
FIGURE 22. A LINEARIZED DEVICE WITH THE 3 RD ORDER TERM ELIMINATED AND A 7.83 CM LENGTH.....	31
FIGURE 23. A DEVICE DESIGN THAT USES A CONSTANT κ AND ABRUPT PHASE CHANGES TO ACHIEVE LINEARIZATION.....	32
FIGURE 24. TRANSFER CURVE FOR THE RECTANGULAR κ -PROFILE OF FIGURE 23 (RED) AND FOR A MACH-ZEHNDER MODULATOR WITH THE SAME SLOPE EFFICIENCY (BLUE).....	32
FIGURE 25. SLOPE EFFICIENCY VERSUS κ . EACH POINT ON THE CURVE IS FOR A DEVICE DESIGN THAT HAS BEEN LINEARIZED. THE OPTIMUM VALUE IS 0.48.....	33
FIGURE 26. DETAILS OF PHASE-REVERSAL TRENCH.....	36
FIGURE 27. DUAL-Y DIRECTIONAL COUPLER MODULATOR.....	37
FIGURE 28. TRANSFER CURVE FOR $s = 3.0$ (BLUE) AND $s = 2.861$ (RED).....	39
FIGURE 29. EVEN AND ODD MODE FORMALISM.....	41
FIGURE 30. MODE-FILTERING PROPERTIES OF A Y-BRANCH WITH A PHASE-SHIFTER.....	42
FIGURE 31. THE ELECTRODE LENGTH, NOT THE COUPLER LENGTH, DETERMINES THE PARAMETERS.....	43
FIGURE 32. AN UNBALANCED Y-JUNCTION.....	44
FIGURE 33. INPUT Y-BRANCH PHASE SENSITIVITY.....	45
FIGURE 34. A DUAL-ELECTRODE SCHEME FOR ADJUSTING THE PHASE AND AMPLITUDE IMBALANCE OF A Y-JUNCTION.....	46
FIGURE 35. A SECOND SCHEME FOR BALANCING THE Y-JUNCTION. THIS USES A STRATEGICALLY-PLACED ELECTRODE (1 OF 4) TO “UNGROW” THE OFFENSIVE ODD MODE.....	46
FIGURE 36. 2 ND HARMONIC REDUCTION AND PHASE REVERSAL USING A COMMON ELECTRODE.....	47
FIGURE 37. THE AMPLITUDE OF IM3 AS A FUNCTION OF DISTANCE ALONG THE COUPLER FOR $s = 3.0$	49
FIGURE 38. SFDR FOR THREE DIFFERENT NOISE BANDWIDTHS (A) 1 HZ, (B) 1 MHZ, AND (C) 1 GHZ AS A FUNCTION OF VELOCITY MISMATCH.....	50

FIGURE 39. SFDR, FOR A 1 HZ BANDWIDTH, AS A FUNCTION OF MODULATION FREQUENCY FOR A DYDCM WITH $s = 2.86059$	52
FIGURE 40. SFDR, FOR A 1 HZ BANDWIDTH, FOR A DYDCM WITH $s = 2.810$	53
FIGURE 41. A Y-FED DIRECTIONAL COUPLER INTENSITY MODULATOR FOR IMDD APPLICATIONS.	54
FIGURE 42. INTENSITY TRANSFER CURVE FOR THE MODULATOR OF FIGURE 41, WITH $s = 2.861$	54
FIGURE 43. A METHOD FOR USING TWO FIBERS, A PHASE-SHIFTER AND A 3-dB COUPLER TO MIMIC THE Y-JUNCTION/PHASE-SHIFTER NEEDED FOR A DYDCM.	55
FIGURE 44. THE ENDS OF THE YFDCM ELECTRODES AFTER ETCHING. THE DEVICES ARE STILL COVERED WITH A THIN LAYER OF RESIST. THE BENDING OF THE ELECTRODES OF THE DEVICE ON THE RIGHT DID NOT CAUSE SHORTING.	56
FIGURE 45. LAUNCH PAD AT THE TOP OF THE MODULATOR. THE RESIST ABOVE THE PADS HAS BEEN REMOVED LITHOGRAPHICALLY SO THAT RF CONTACT CAN BE MADE.	57
FIGURE 46. MEASURED TRANSFER CURVE FOR A YFDCM. THE TWO COLORS ARE FOR THE TWO DIFFERENT OUTPUT PORTS.	58
FIGURE 47. HARMONIC MEASUREMENT SYSTEM.	59
FIGURE 48. THE 1 ST , 2 ND AND 3 RD HARMONIC AS A FUNCTION OF DC BIAS VOLTAGE ON THE MODULATOR.	60
FIGURE 49. 3 RD HARMONIC POWER VS DRIVE LEVEL FOR DEVICE Y31G4B, WAFER 750N4.	61
FIGURE 50. 3 RD HARMONIC POWER VS DRIVE LEVEL FOR DEVICE Y30G4B, WAFER 750N4.	61
FIGURE 51. THE ELECTRIC FIELDS IN A LiNbO ₃ MODULATOR WHEN BIAS DRIFT IS A PROBLEM. WITH TIME, THE DASHED FIELD LINES DISAPPEAR, AND ALL OF THE FIELD ENDS UP IN THE SiO ₂ BUFFER LAYER.	62
FIGURE 52. HARMONIC DATA FOR DEVICE Y33G2A FROM WAFER 800D. THIS WAFER HAS THE ANTI-DRIFT COATING.	63
FIGURE 53. 3 RD HARMONIC DATA FOR THE DEVICE OF FIGURE 52.	63
FIGURE 54. THE DEVICE OF FIGURE 53, MEASURED A SECOND TIME (THE RED CURVE).	64
FIGURE 55. 3 RD HARMONIC POWER AS A FUNCTION OF DEVICE TEMPERATURE FOR A FIXED DRIVE LEVEL.	64
FIGURE 56. A FABRY-PEROT ETALON. THE SMALL RED ARROWS INSIDE REPRESENT THE LIGHT REFLECTED AT EACH SURFACE.	65
FIGURE 57. THE TRANSMITTANCE OF A LiNbO ₃ /GLASS ETALON AS A FUNCTION OF OPTICAL LENGTH.	65
FIGURE 58. MEASURED 3 RD HARMONIC DATA (POINTS) AND THE FP MODULATOR 3 RD HARMONIC POWER (RED LINE) FOR TWO DEVICES FROM THE SAME WAFER. THE BLUE AND BLACK POINTS IN THE LEFT FIGURE ARE FOR MEASUREMENTS MADE AT TWO DIFFERENT TIMES THE BLUE CURVE IN THE RIGHT IS THE PREDICTED RESPONSE FOR A DEVICE WITH NO FP EFFECT.	66
FIGURE 59. 3 RD HARMONIC POWER FOR DEVICE Y34G3B, WAFER 800D3. THE BLACK DATA POINTS WERE TAKEN AFTER A 6° ANGLE HAD BEEN FORMED AT THE END OF THE WAFER, THE RED DATA TAKEN BEFORE.	67
FIGURE 60. A PLOT SHOWING HOW SLOPE-3 EFFECTS WILL ALWAYS MANIFEST THEMSELVES WHEN THE PRIMARY DISTORTIONAL TERM IS SLOPE 5.	68
FIGURE 61. 3 RD HARMONIC POWER FOR DEVICE Y33G2A, WAFER 800D2, CORRECTED TO REMOVE THE FABRY-PEROT MODULATOR EFFECT.	68
FIGURE 62. THE UP-CONVERSION OF BASEBAND NOISE TO “SKIRT” NOISE DUE TO MODULATION AT F_c	73

LIST OF TABLES

TABLE 1. TOLERABLE PHASE LOCKING ERROR VERSUS MODULATOR EXTINCTION RATIO AND NOISE BANDWIDTH.	8
TABLE 2. SFDR CONFIGURATIONAL HANDICAPS.	18
TABLE 3. SLOPE EFFICIENCIES FOR VARIOUS CONFIGURATIONS.	34
TABLE 4. MANUFACTURING TOLERANCES.	35
TABLE 5. SFDR VERSUS MEASUREMENT BANDWIDTH FOR TWO DIFFERENT VALUES OF S	69

EXECUTIVE SUMMARY

The goal of this program was to demonstrate an optical link having a Spur Free Dynamic Range (SFDR) in excess of 130 dB in a 1-Hz bandwidth, at 10 GHz. The approach chosen to do this was to use coherent amplitude modulation (AM) to achieve a high initial value of SFDR ($> 120 \text{ dB Hz}^{2/3}$), and to then develop a linearized modulator that would raise this respectable number to $> 130 \text{ dB Hz}^{4/5}$.

The program was divided into two parts. In the first part a coherent AM optical link was developed that demonstrated a SFDR of $120 \text{ dB Hz}^{2/3}$ using a 100 mW laser and Commercial Off-The-Shelf (COTS) modulators and photodetectors. A slightly higher value ($123 \text{ dB Hz}^{2/3}$) was obtained when the laser power was boosted to 500 mW with a fiber optical amplifier. The link, which used polarization-maintaining components and fiber throughout, and fusion-splices instead of connectors, had none of the instabilities historically associated with homodyne systems. The system locked, and held lock, for the duration of all measurements (which often lasted all day). We saw no damage effects in the modulators used, even at the 500 mW power levels used at the program's end.

The plan to raise the $120\text{-}123 \text{ dB Hz}^{2/3}$ SFDR values demonstrated on the first part of this program to $130 \text{ dB Hz}^{4/5}$ by using a linearized directional coupler modulator was pursued in the second part of the program, but was not completed. Y-fed directional coupler modulators having the correct configuration were loaned to us by the PACT* program office, and an extensive set of measurements was conducted on these devices. After resolving a very serious measurement problem, it was found that the devices did indeed perform as theoretically predicted at low frequencies. The devices were stable, did not drift once equilibrium was reached, and showed the linearization predicted by theory. A logarithmic slope of 4.6 was demonstrated for the distortional components of the best device; the only reason a perfect slope-5 was not demonstrated was because the 3.3-cm device tested was 66 microns too long, and not because it had any inherent flaws.

Our analysis did show, however, that Y-fed linearized directional coupler modulators (and perhaps all linearized directional coupler modulators) have a shortcoming that makes them unsuitable for wideband operation, at least for coherent AM applications: RF attenuation in the Co-Planar Strip (CPS) traveling-wave electrodes degrades the linearization at higher frequencies. The present devices can be designed to operate at a specific center frequency with maximum performance, but the linearization bandwidth at this frequency is less than 2 GHz.

Our measurements on the first part of the program also suggested that the photodetectors were beginning to limit the SFDR of the unlinearized link. Although this will have to be looked into more rigorously before any definitive statement can be made, the possibility exists that today's photodetectors could limit the performance of optical AM links to $\sim 127 \text{ dB Hz}^{2/3}$.

*Photonic A to D Converter Technology, AFRL Rome/DARPA-MTO

1.0 Introduction

Fiber optic links are well known for their ability to transmit very large amounts of information over long distances. However, they have historically suffered from a Spur Free Dynamic Range (SFDR) that is about 20 dB lower than that achievable with conventional microwave links. Although this shortcoming is not a handicap for those systems that transmit information in a digital format (where distortion is not a real issue), it can be debilitating for high-performance analog systems, especially those that must process both weak and strong signals simultaneously.

The root of the problem for intensity modulation/direct detection (IMDD) links is the nonlinearity of the modulator. The distortion generated by all other components is either small (as is the case for the photodiode at modest photocurrents [1]), or only plays a role when the drive voltage required by the modulator (V_π) is high (so that preamplifier distortion becomes an issue [2]). Analyses have shown that a better optical modulator would, in fact, allow one to equal or perhaps even surpass the performance of microwave links.

In view of this, a world-wide quest for a highly linear optical modulator was started back in the early nineties. The results have been a large number of theoretical schemes to linearize modulators, most of which require extremely tight fabrication tolerances to achieve large increases in SFDR. In almost all of these schemes, one linearizes the modulator by generating a secondary distortional component that cancels the distortion generated by the primary modulator. The improvement in SFDR is thus directly related to how well one can subtract one large distortional term from another. If the cancellation is not complete, the actual improvement in SFDR can be considerably less than the theoretical predictions. Thus, while appearing attractive on paper, almost all of these schemes suffer degradation in implementation, generating at best 5 to 10 dB of improvement in SFDR at larger bandwidths.

A SFDR improvement of 5 to 10 dB, however, would be enough if the SFDR of the unlinearized link were already 120 to 125 dB $\text{Hz}^{2/3}$. And coherent AM links, when properly configured, can attain such values. The two goals for this program were to first demonstrate a SFDR of greater than 120 dB $\text{Hz}^{2/3}$ for an unlinearized AM link, and to then enhance this value by using a linearized directional coupler modulator to achieve the *small amount* of linearization needed to achieve a SFDR of 130 dB. We succeeded in accomplishing the first of these goals, but were not entirely successful in achieving the second.

2.0 Coherent AM Links

2.1 Overview

Earlier studies showed that coherent AM links had an intrinsic 6 dB SFDR advantage over Intensity-Modulated/Direct-Detection (IMDD) links [2]. What was not realized at the time was that one can operate a coherent AM link in a manner that would achieve a much greater

enhancement in SFDR. The concept is quite simple, but to understand it, one must first appreciate the difference between Intensity and Amplitude modulation.

Figure 1 shows the transmission profiles for the optical *intensity* (Figure 1b) and the optical *amplitude* (Figure 1c) for a conventional interferometric modulator. Note that biasing the modulator for minimum intensity also gives minimum amplitude. However, note also that operation about this bias point gives a linear response in amplitude, (but a quadratic response in intensity). A coherent AM link is thus intentionally operated at this linear pinch-off point, which means that no light is transmitted by the modulator when the modulation signal is zero.

When biased at pinch-off there is a certain flexibility in how one generates the output signal. For example, if 1 milliwatt of average transmitted power is desired, one could use a laser with a few milliwatts of CW power, and then drive the modulator with a large modulation signal. Conversely, one could employ a laser with 100 milliwatts of optical power, and drive the modulator with a weaker modulation signal. The advantage of performing the latter - rather than the former - is that one is operating in a much more linear regime of the modulator. Because the time-averaged output power is the same for both cases, the power seen by the receiver will be identical. Therefore, operating the system with a higher input power (P_{in}^0 in Figure 1), but with a weaker modulation signal to enhance linearity, has no detrimental impact on the rest of the system.

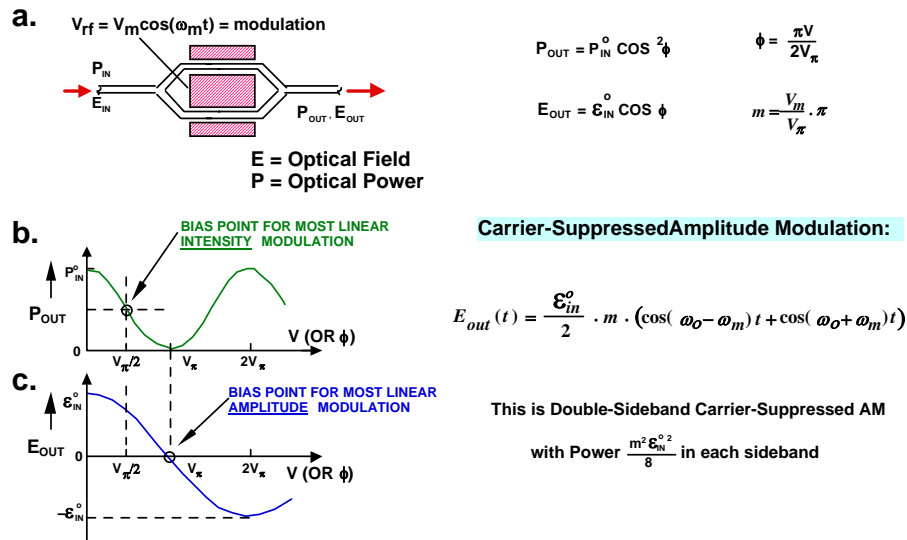


Figure 1. Amplitude versus Intensity modulation.

In fact, the received optical power is quite small. An AM system operated at pinch-off has a fully-suppressed carrier, so that only the sideband energy reaches the photodetector. Because the maximum sideband energy allowed by the SFDR condition is a very small fraction ($< 0.5\%$) of the total laser energy, the received power for a 100 milliwatt laser will be a fraction of a milliwatt, which is substantially less than the local oscillator power (~ 2 milliwatts). The AM receiver power-handling requirement is thus very modest, even for 100 mW laser powers, and well within the capability of today's photodiodes [1].

2.2 Theoretical Performance of an Unlinearized AM Link

The Spur Free Dynamic Range (SFDR) is the ratio, expressed in dB, of the largest undistorted signal to the smallest signal measurable in an analog microwave system. The largest signal is defined to be one for which the biggest spur is just equal to the noise floor, the smallest signal is defined to be one that is just equal to the noise floor. The SFDR is therefore the maximum Signal to Noise Ratio (SNR) consistent with distortionless operation of a system.

Figure 2 shows a simple AM system. The Mach Zehnder modulator is biased at V_π so that the carrier is completely suppressed. The modulated light and local oscillator (LO) laser light are combined and mixed on a photodetector (an optical squarer). When properly phase-locked, the system recovers the information signal, $\theta(t)$. For a single RF tone, $\theta(t)$ is given by $\theta_o \cos(\omega_m t)$.

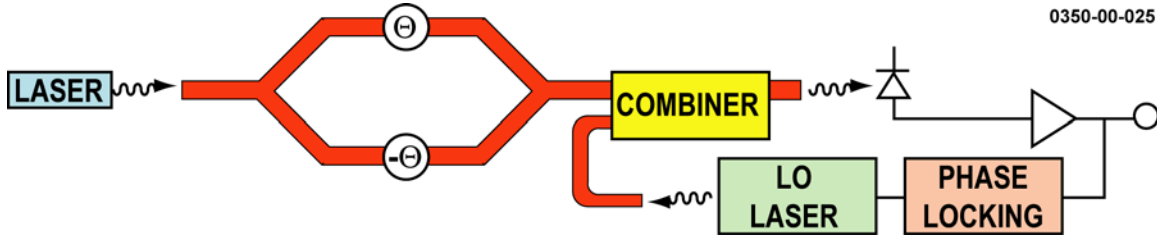


Figure 2. A simple AM link using a Mach Zehnder modulator and homodyne detection.

To follow this mathematically, one starts with a modulated signal. The output of the Mach Zehnder is given by [2]

$$E_s(t) = E_s \cos(\omega_c t) \sin(\theta(t)) \quad (1)$$

When θ is small, this can be approximated by

$$E_s(t) = E_s \cos(\omega_c t) \theta(t) \quad (2)$$

Equation (2) is an AM signal with fully suppressed carrier; this is the form the signal *must* have for the approach that we are proposing, regardless of the type of modulator used. The local oscillator signal is

$$E_{lo}(t) = E_{lo} \cos(\omega_c t + \phi) \quad (3)$$

When these two lightwaves are added and squared at the photodetector, and averaged over several optical cycles, one finds that

$$\left\langle \left(E_s \cos(\omega_c t) \theta(t) + E_{lo} \cos(\omega_c t + \phi) \right)^2 \right\rangle = \frac{1}{2} E_s^2 \theta^2 + \frac{1}{2} E_{lo}^2 + 2 E_s E_{lo} \theta \langle \cos(\omega_c t) \cos(\omega_c t + \phi) \rangle \quad (4)$$

Reducing the cosine product to the cosines of the sum and difference of the arguments, and performing the time average gives the photodetector current, I .

$$I = A \left(\frac{1}{2} E_s^2 \theta^2 + \frac{1}{2} E_{lo}^2 + \theta E_s E_{lo} \cos(\phi) \right) \quad (5)$$

where A is a constant that converts the optical fields, squared, into an electrical current.

The purpose of the locking circuit in Figure 2 is to adjust and hold the angle ϕ at the value that gives maximum recovery of the modulation information. For this configuration, that angle would be any integer multiple of π . Setting $\cos \phi$ equal to 1, and recognizing that $\frac{1}{2} A E_s^2$ and $\frac{1}{2} A E_{lo}^2$ are the DC photocurrents for the source and local oscillator laser, respectively, one can rewrite this as

$$I = I_s \theta^2 + I_{lo} + 2\theta \sqrt{I_s I_{lo}} \quad (6)$$

In computing the SFDR for weak received signals, one can normally ignore the first term (more about this later).

The dominant source of distortion in this link is the nonlinearity of the $\sin(\theta)$ term in eqn (1). Using Fourier-Bessel expansions, one can show that the power in the third order intermod (IM₃) divided by the power in one of the fundamentals for two-tone excitation, each of amplitude θ_o , is given by [2]

$$\frac{P_{IM_3}}{P_1} = \left(\frac{J_2(\theta_o)}{J_o(\theta_o)} \right)^2 \cong \left(\frac{\theta_o^2}{8} \right)^2 \quad (7)$$

The electrical signal power at the detector for one of the tones is given by the third term of equation (6),

$$P_1 = 4 I_s I_{lo} \langle \theta^2(t) \rangle = 2 I_s I_{lo} \theta_o^2 \quad (8)$$

The maximum power occurs when the spur power just equals the noise power. For a reasonably high local oscillator power, ($I_{lo} > 2$ mA), one can ignore the kT noise of the detector amplifier. The noise is then just the photonic shot noise plus the Relative Excess Noise (REN) of the laser (Relative Intensity Noise minus the shot noise).

$$P_1(\theta_o) \left(\frac{\theta_o^2}{8} \right)^2 = P_{noise} = 2e I_{lo} B + I_{lo}^2 REN B \quad (9)$$

where B is the noise bandwidth. Substituting equation (8) into (9) gives the sixth power of θ_o which can be easily solved to give $P_{1 \max}$. One then computes the SNR using this value, finding

$$SFDR = 4 \left(\frac{I_s}{eB + \frac{1}{2} I_{lo} REN B} \right)^{2/3} \quad (10)$$

For a laser equivalent current of 100 mA and zero REN, the SFDR for a 1 Hz bandwidth is 124.6 dB. Reducing this to 50 mA gives a SFDR of 122.6 dB.

I_s in this equation is the photodetector current that one would measure with the local oscillator laser turned off, and with the modulator biased for maximum throughput. (One doesn't really do this; this is just a definition). Thus, if the modulator had excess loss (as they all do), the laser power would have to be raised to compensate for this loss.

There are three shortcomings of the detection scheme shown in Figure 2 that we can easily remedy. First, our derivation assumed 100% efficient combining of the light. However, combiners have a maximum theoretical efficiency of only 50 %. Thus half the LO and signal light are lost, which reduces the SFDR. Second, in deriving this relation, we purposely neglected the first term in equation (6). When I_s is small, this is justified. However, for our application, I_s is quite large, which means that this first term will generate a non-negligible 2nd harmonic. And third, performance is degraded by REN noise in the photodetector, especially if one raises the LO power to increase link gain.

These three shortcomings can be eliminated by using the dual-balanced detector shown in Figure 3.

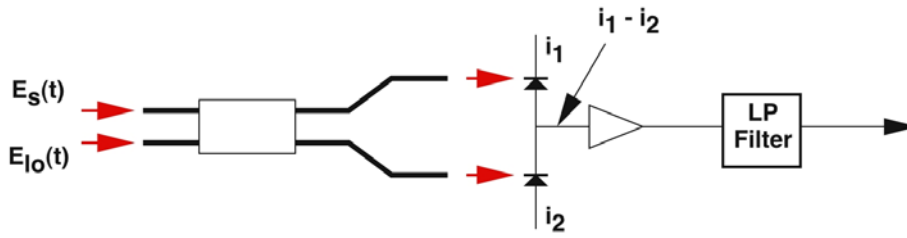


Figure 3. A dual-balanced detector.

The two photodiodes are connected in a way that gives high common mode rejection. For common mode illumination (i.e., equal signals on both detectors), all of the current generated in the top photodiode flows through the bottom photodiode, and none flows into or out of the transimpedance amplifier. It is trivial to show that the current into the amplifier is given by the differential signal $I_1 - I_2$, where I_1 and I_2 are the top and bottom photocurrents, respectively. Because of this feature, and the symmetric split of the directional coupler, fluctuations in the local oscillator power do not show up at the amplifier input. Nor do the time variations generated by the first term in equation (6), which could generate undesirable second harmonics. Furthermore, because all of the energy from both lasers reaches the photodiodes, the process is 100 % efficient.

To see better how this works, consider again the signal and local oscillator fields given by equations (2) and (3). Light passing through the 3 dB coupler is split and sent to the two output ports, with no phase shift for the thru port, and a 90° shift for the cross port. Thus, the photocurrent at the top detector will be given by

$$I_1 = \frac{1}{2} \left\langle \left(E_s \cos(\omega_c t) \theta(t) + E_{lo} \sin(\omega_c t + \phi) \right)^2 \right\rangle = \frac{E_s^2 \theta^2}{4} + \frac{E_{lo}^2}{4} + \frac{E_s E_{lo} \theta \sin \phi}{2} \quad (11)$$

and that at the bottom by

$$I_2 = \frac{1}{2} \left\langle \left(E_s \sin(\omega_c t) \theta(t) + E_{lo} \cos(\omega_c t + \phi) \right)^2 \right\rangle = \frac{E_s^2 \theta^2}{4} + \frac{E_{lo}^2}{4} - \frac{E_s E_{lo} \theta \sin \phi}{2} \quad (12)$$

The current flowing into the amplifier is given by

$$I_1 - I_2 = E_s E_{lo} \theta(t) \sin \phi \quad (13)$$

which is the same as the time-varying term in equation (6) except for a 90° shift in ϕ (produced by the 90° hybrid). This differential signal is maximized when ϕ is 90° . Note that the currents that would produce the REN and the term that would produce a 2nd harmonic are gone. The photonic shot noise at each detector, however, is uncorrelated, and so doesn't subtract. Thus the REN noise term disappears from equation (9), but the shot noise term does not. The equation for the SFDR of a system using a balanced detector is therefore

$$SFDR = 4 \left(\frac{I_s}{eB} \right)^{2/3} \quad (14)$$

2.3 Link Gain and Noise Figure

The SFDR is the most important figure of merit for any link. The link gain and noise figure, however, are quantities often of more interest to system designers. The system gain is given by

$$G = 4 I_{lo} I_s \left(\frac{\pi}{2 V_\pi} \right)^2 Z_o^2 \quad (15)$$

and the noise figure by

$$NF = 1 + \frac{1}{2 I_s \left(\frac{\pi}{V_\pi} \right)^2 Z_o \frac{kT}{e}} \quad (16)$$

where Z_o is the characteristic impedance of the system (typically 50 Ohms) and V_π is the modulator switching voltage. Equation (16) assumes that the kT noise following the photodetector is negligible compared to the shot noise. If the photodetector has an internal 50-Ohm resistor (as many do), the link gain will be reduced by 6 dB; the noise figure, however, will remain unchanged.

2.4 Laser Noise

The fact that the dual balanced detector removes the REN term from equation (10) does not mean that the effects of REN have disappeared altogether. Because the detected signal is proportional to the product of the amplitudes E_s and E_{lo} (see equation (13)), fluctuations in the amplitude of both the source and local oscillator lasers can affect the received signal. We have analyzed this process in some detail in Appendix I, and have found that laser fluctuations can add noise skirts to the received signal. These noise skirts will be narrow (< 1 MHz) and should not affect the SFDR unless one is measuring tones that are within 1 kHz of each other. We refer the interested reader to the appendix for more details.

2.5 Finite Extinction

Mach Zehnder modulators separate the light at the input, run it through two arms, and then recombine it after having shifted the phase in each arm. If the amplitudes at the point of recombination are equal, the extinction ratio (the on/off ratio) can be virtually infinite. However, if one arm of the interferometer attenuates the light just slightly more than the other, or if the polarizations rotate ever so slightly, then the two fields cannot subtract completely, and the ratio will be finite. Although the power leaking through the device will normally be quite low, the field amplitudes that correspond to this power can be significant. For example, an extinction ratio of 20 dB means that the emerging field has an amplitude that is 10% of E_s . Commercial devices have extinction ratios that range between 20 and 30 dB.

We have analyzed this problem in Appendix II, and have found that two things must go awry before there is any degradation in SFDR: (1) the extinction ratio must be finite, and (2) there must be a phase locking error in the receiver. Poor extinction by itself does not cause a problem. One can show that the complete signal is recovered, with no adverse effects. However, if one has finite extinction, and the phase of the LO is not some exact multiple of $\pi/2$, one sees a small second harmonic.

A useful number is the phase tolerance that would give a 2nd harmonic SFDR equal to the 3rd order intermod SFDR. For a 1 MHz bandwidth and an extinction ratio of 20 dB (the worst value that one would expect), the phase error must be less than 6.8 degrees, which is a reasonable number to achieve. Table 1 shows the phase locking tolerances for several bandwidths and extinction ratios. These numbers define the accuracy with which one must lock the receiver to avoid a 2nd harmonic.

Table 1. Tolerable phase locking error versus modulator extinction ratio and noise bandwidth.

Bandwidth	$\Delta\phi$ (20 dB)	$\Delta\phi$ (30 dB)
1 Hz	0.68°	2.17°
1 MHz	6.8°	21.7°
1 GHz	21.6°	63.7°

2.6 The Use of an Optical Amplifier to Increase the SFDR

Equation (14) shows that the SFDR increases as $I_s^{2/3}$, where I_s is the equivalent detector current of the light that would reach the photodetectors if the modulator were biased for maximum transmittance. An I_s of 100 mA, for example, gives a SFDR of $124.6 \text{ Hz}^{2/3}$, an excellent value for an unlinearized system. However, achieving such a value is not easy. The insertion loss of the modulator (typically 3 - 4 dB), the less-than-ideal quantum efficiency of the photodetectors (0.75 Ampere/Watt), the power that must be tapped off to provide the local oscillator signal, and the small but non-negligible insertion loss of other optical components all conspire to drive up the value of the actual laser power required. To get an I_s of 100 mA, one typically needs greater than 300 mW of laser power.

Single-line narrow-linewidth lasers that produce $> 300 \text{ mW}$ of power at $1.55 \mu\text{m}$ are not yet available commercially (though they may be soon). However, one can use an optical amplifier to raise the power to almost any value desired. The question is whether this can be done in a way that increases the SFDR performance of the link. The answer is that it can, if correctly implemented.

The fundamental noise associated with any optical link is the photonic shot noise. This noise is related to the statistical nature with which photons are emitted by the laser. If one generates an average of 10 photons in one second, the rms fluctuation in that number is $\sqrt{10}$, or about 3.2. This rms deviation is the noise; hence, the SNR is $10/3.2$, or 3.12. If one increases the number of photons per second to 1000, the deviation is 32, so the SNR is 1000 divided by 32, which is 31.2. Thus, by simply increasing the photon flow rate, one increases the SNR.

Now, suppose one amplifies this weak photon source (10 photons \pm 3.2 per second) with a noiseless optical amplifier having a gain of 100. The optical beam now has 1000 \pm 320 photons per second. Thus, even though the photon number has increased markedly, the SNR ($1000/320 = 3.12$) has not. The fluctuations of the original source are simply multiplied by the gain of the amplifier.

This, of course, is how it has to be. An amplifier cannot improve the SNR of an incoming signal; it can only make it the same or worse. Thus, a MOPA (Master Oscillator/Power Amplifier) consisting of a 10 mW source laser, and having a gain of 20 dB, will produce a 1 Watt beam having the SNR of a 10 mW beam. The SNR is thus set by the source laser, not the following amplifier.

Why, then, should one use a MOPA? If one were to receive all of the power generated by this system, one would have a SNR no better than that of the weaker source laser, so why bother? The reason one bothers is that in a very large number of applications, not only does one not receive all of the power, but more often than not, one receives only a very small fraction of the broadcast power. Communication and signal processing systems routinely divide a high-power signal into hundreds of much weaker signals, or propagate it through space where a large fraction of the power is lost to beam divergence (Lasercom systems). Why are MOPAs useful in these applications?

The reason is simple, and can best be understood by the Lasercom example. Here one typically has a source of 10 milliwatts that is amplified to 1 Watt. The SNR of this broadcast signal is that of a 10 milliwatt source. However, due to beam spreading, the received beam at a one-geosynch distance is about 1 microwatt! The SNR of this signal is thus that of a 1 microwatt laser, which is appreciably worse than that of a 10 milliwatt laser. Thus, it doesn't matter that the broadcast beam has the SNR of a 10 milliwatt source, and not that of a 1 Watt source, because the statistics of the received beam are that of a 1 microwatt source, which is 40 dB lower than that of the 10 milliwatt source.

A more complete noise analysis, which includes the additive noise of the optical amplifier, shows that MOPAs are useful in those applications where the product of the MOPA amplifier gain, G , multiplied by the system attenuation that *follows* the amplifier, A , is substantially less than unity (Figure 4). Thus, if $GA \ll 1$, the MOPA amplifier does not corrupt the system SNR.

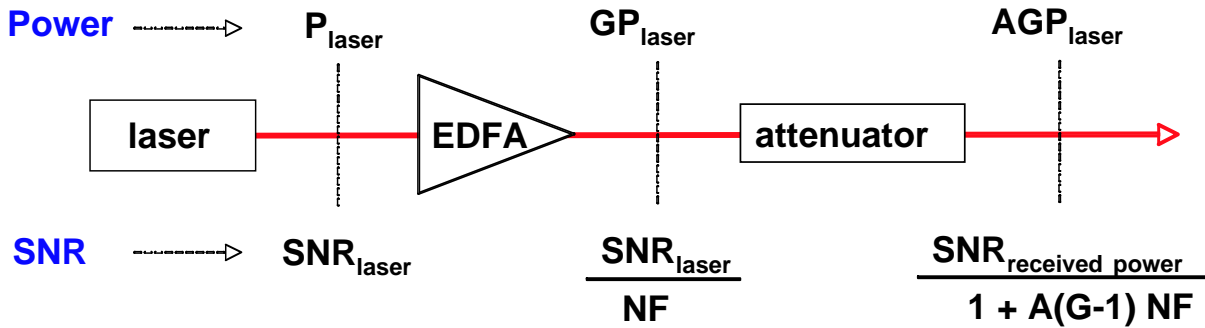


Figure 4. SNR versus position in link for a system using an optical amplifier of gain G , followed by an attenuator having a value of A .

The source of attenuation in our AM link is the carrier-suppressed mode of operation: the modulator is biased at pinch-off so that with no modulation, no power reaches the photodiode. When a modulation signal *is* applied to the modulator, the only power that reaches the photodetector is that of the generated sidebands, which are very weak (a few tenths of a milliwatt). Thus, the modulator attenuates the parent source by almost 30 dB. Hence, for our case where the desired gain (G) is 3 (100 mW \rightarrow 300 mW), the condition that GA is $\ll 1$ is well-satisfied ($3 \times .001 = .003$, which is $\ll 1$). The fact that we are dealing with homodyne detection in a coherent system does not invalidate this conclusion: the fundamental $GA \ll 1$ rule still applies.

3.0 SFDR Measurements for the Unlinearized AM link

3.1 Experimental Setup

A schematic diagram of the link in its simplest form is shown in Figure 5. The link consists of a medium-power laser, a directional coupler that splits off a small amount of laser power for use in the homodyne mixing process (i.e., for use as a local oscillator signal), an electrooptic modulator, a 50:50 4-port directional coupler, a dual-balanced photodetector, and the locking circuit needed for stable operation.

The locking circuit ensures that the phase of the local oscillator and signal lightwaves are always in quadrature. If they are not, i.e., if the relative phase of either lightwave is shifted, the signal will be reduced; if it is shifted by 90° , it goes to zero (see, for example, equation (13)). Therefore, in order to maximize the received signal and to eliminate slow fluctuations in that signal, one must hold the differential phase between the two arms of the interferometer to a fixed value (an odd-integer multiple of 90°). The locking circuit shown in Figure 6 does this by using a dithering technique to adjust the length of one branch of the fiber interferometer so that the RF signal picked off at the 20 dB directional coupler is maximized. This particular technique was found to be superior to another technique that monitored the DC photocurrent at either detector. The only drawback was the need for additional components: an RF coupler, an amplifier and a diode power detector.

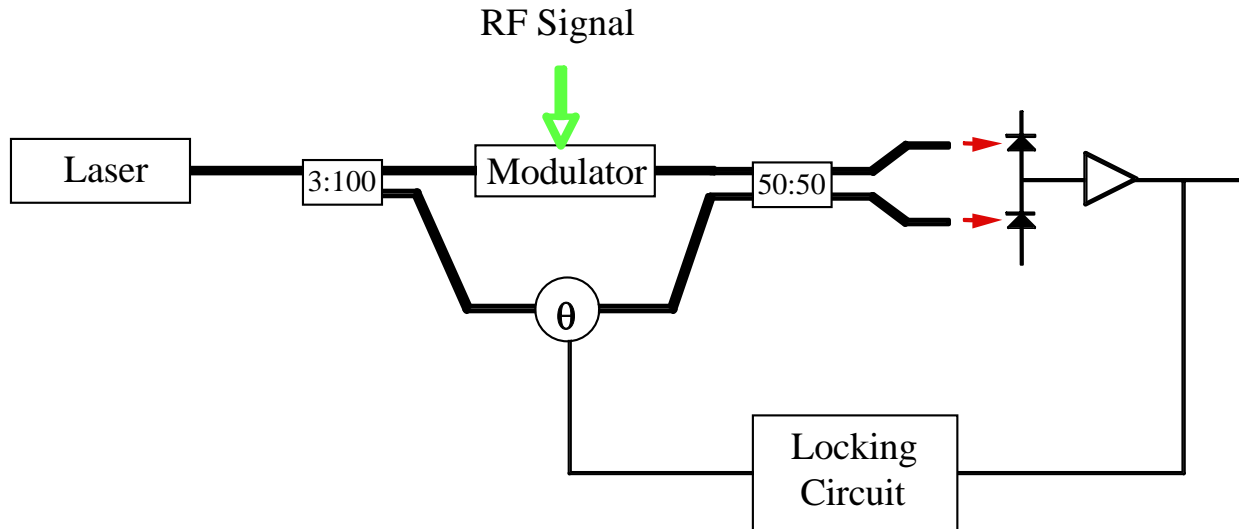


Figure 5. A simple homodyne AM link.

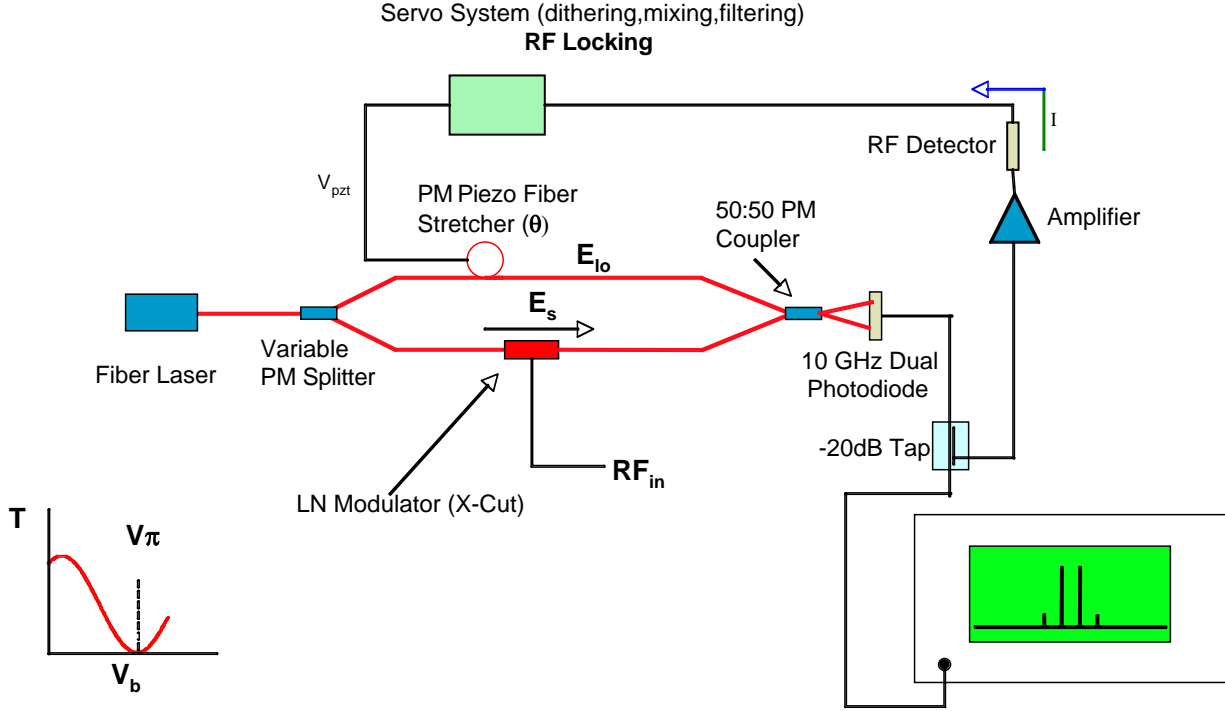


Figure 6. The experimental link with details of the locking technique shown.

A photo of the experimental layout and the ancillary electronics is shown in Figure 7. The unit in the center of the picture is the spectrum analyzer used to measure the fundamental signal and the 3rd order intermodulation products, the two instruments to the lower right are the signal sources, the small black box sitting on top of the instruments is the optical amplifier used in an experiment that we will discuss later, and the box under it is that amplifier's power supply. Pictures of individual components of interest are the Mach-Zehnder modulator (Figure 8), the piezoelectric fiber stretcher that was used in the control circuit to mechanically change the length of one of the fibers (Figure 9), the dual-balanced photodetector (Figure 10), and an older picture of the experimental setup that shows the commercial locking electronics (the blue-yellow box) used to drive the line stretcher (Figure 11).

With the exception of the dual-balanced detector and the optical amplifier used later, all components in the link were Polarization-Maintaining (PM). The Mach-Zehnder modulator was special-ordered to have PM fibers on the input and output, the directional couplers and variable splitter were made with PM fiber, and the source laser also had a PM output fiber. Furthermore, with two exceptions (the laser and the dual-balanced photodetector), all connections were fusion-spliced.

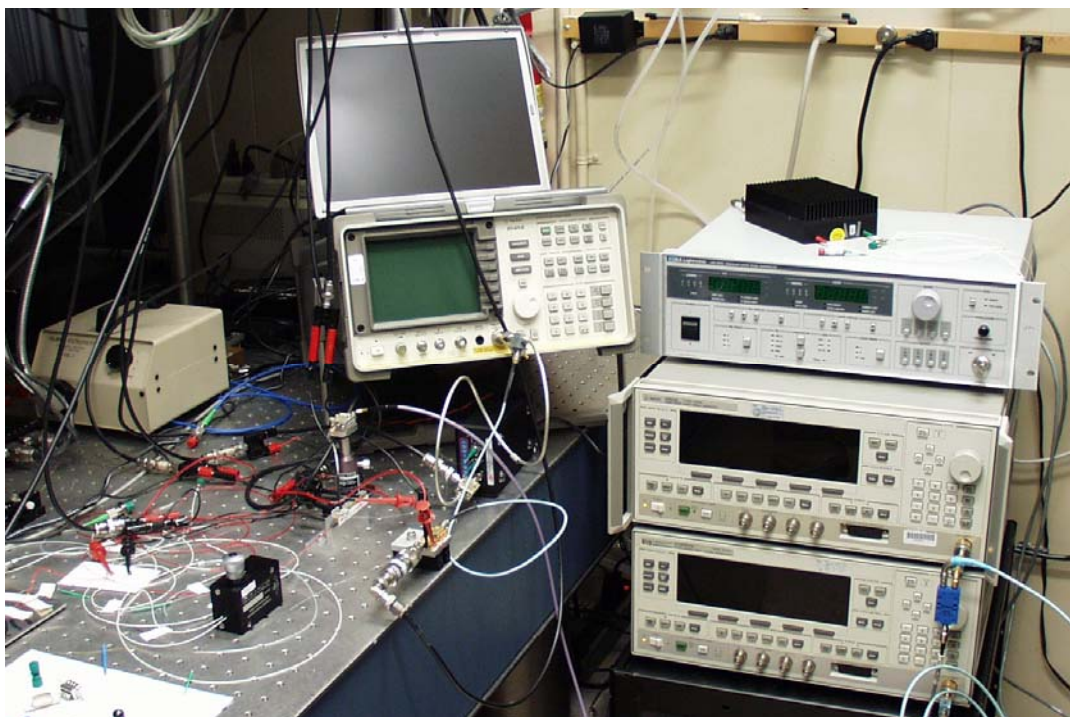


Figure 7. Experimental setup and support equipment.

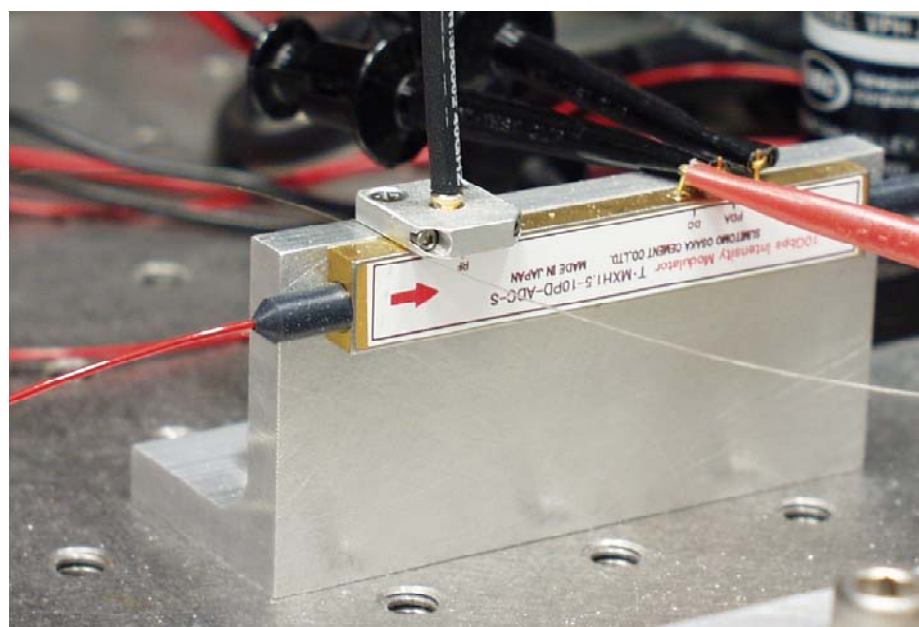


Figure 8. The X-cut (push-pull) Mach-Zehnder modulator.

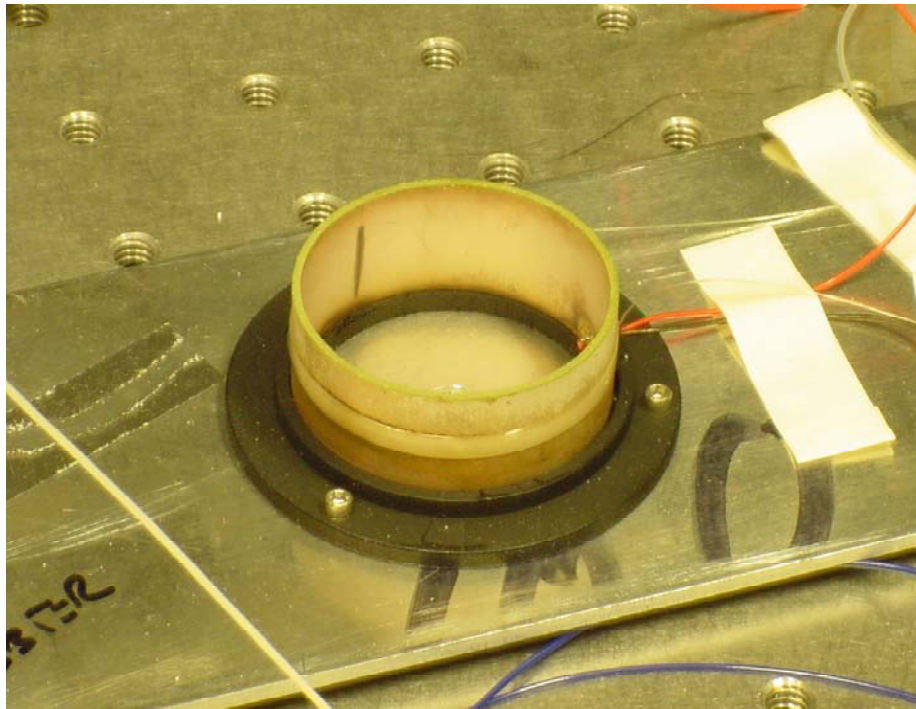


Figure 9. Piezoelectric line stretcher wound with PM fiber.

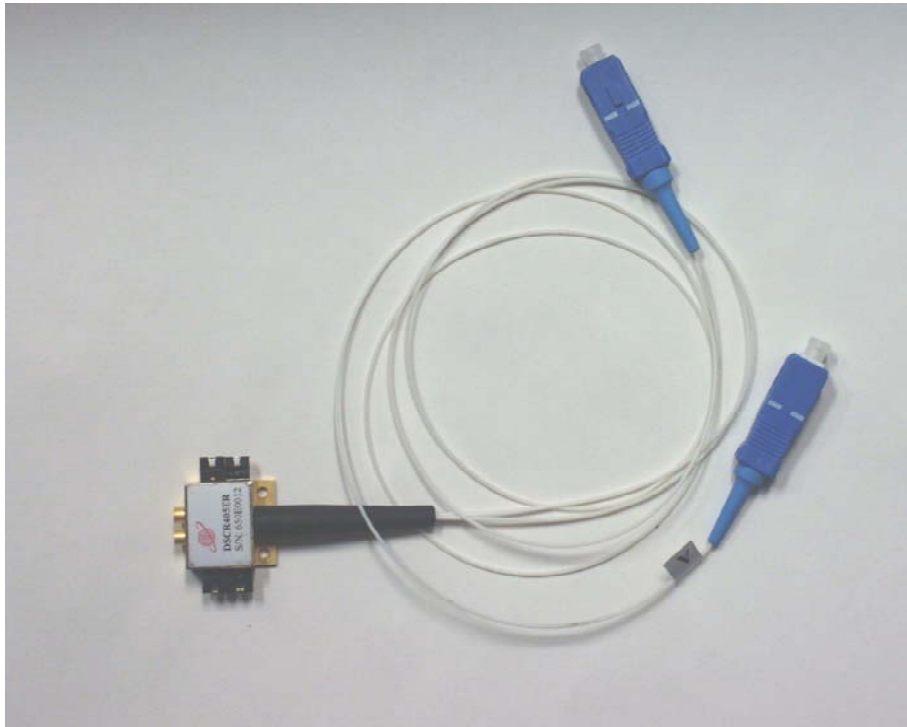


Figure 10. The dual-balanced photodetector.

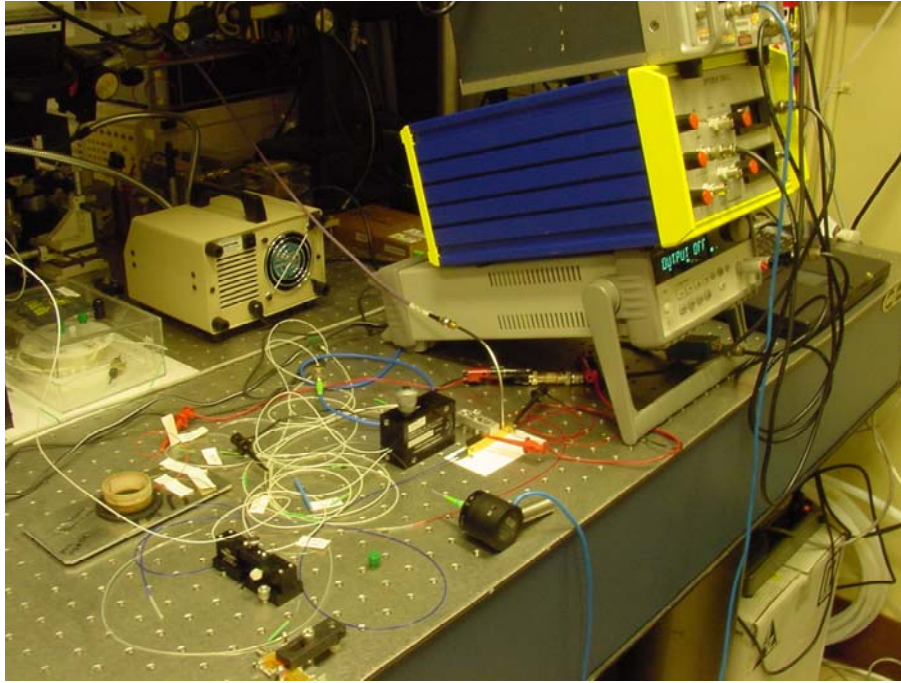


Figure 11. An earlier setup showing the control electronics (the blue-yellow box).

3.2 Measurement Technique

The SFDR was measured using the standard two-tone approach, in which two tones, separated in frequency by a small amount (typically 80 MHz for the measurements reported here), are simultaneously applied to the input of the link. Initial measurements were performed at a 1 GHz center frequency, but as better equipment became available and as our confidence in the measurement technique increased, we progressively raised the center frequency to 3 GHz and then to 9-10 GHz. There was no measurable difference in the SFDR at the different frequencies, which is what one would expect for a link in which the modulator distortion dominates. All final measurements were performed at 9 or 10 GHz.

A recurring problem with measurements of this sort is interference from undesired intermodulation tones generated by nonlinearities in the measurement equipment. The signal sources themselves, the spectrum analyzer used to measure all tones, and any microwave amplifiers used in the circuit, either as preamplifiers or as postamplifiers, can all produce spurs that mask those produced by the optical link. A crucial requirement for an accurate measurement of the link SFDR, therefore, is that the spurious intermodulation tones generated by the ancillary equipment be negligible compared to those generated by the link.

To eliminate spurs caused by the signal sources, microwave isolators were used at the sources' output in order to prevent nonlinear interactions between the two. A tunable YIG filter was used to suppress spurs generated whenever a microwave amplifier was used between the dual-balanced photodetectors and the spectrum analyzer. Spurs in the spectrum analyzer were identified by changing the values of the electronically-switchable attenuators within the analyzer, and monitoring the change in the measured intermodulation tone: if a 10 dB change in the attenuation changed the spur in question by 10 dB, then the analyzer was not producing an interfering spur; if the change was greater than 10 dB, then a higher level of input attenuation had to be used.

Measurements at the beginning of the program were made by first measuring the fundamental and intermodulation tones directly, (i.e., without the use of a postamplifier), and then measuring the noise by installing a high-gain, low-noise postamplifier to raise the noise well above the noise floor of the spectrum analyzer. This technique, while simple and fundamentally correct, had the drawback that it required an accurate measurement of the amplifier gain at the fundamental tone and intermod frequencies.

Toward the end of the program, a postamplifier with a high IP3 was used, together with a YIG filter, to measure the fundamental and intermodulation tones and the link noise. This technique did not require any measurement of the postamplifier gain. It also allowed one to measure the noise in the presence of a large signal, a necessary feature for links having upconversion noise (a mixing between low-frequency noise and the fundamental tones). We observed no up-conversion effects for our link.

3.3 Configurational Effects

One of the more interesting effects that we found was that the SFDR one measures depends on the geometrical configuration of the modulator and/or detector used. By this we mean whether the modulator has a single-sided or push-pull drive, or whether the detector is single or dual-balanced. The dependence for the detector was expected, that for the modulator was not.

Figure 12 shows a Mach-Zehnder z-cut modulator with single-sided drive. One sees that the hot electrode is above one guide only. Because the field under this electrode is 7-10 times stronger than that under the ground plane electrode, the refractive index change for the left waveguide is 7-10 times greater than that for the right. The modulator effectively has a single-sided drive because the lion's share of the phase change is occurring in the left waveguide, with almost none occurring in the right.

The push-pull modulator shown in Figure 13, on the other hand, has waveguides situated to the left and right of the hot electrode. Because the index change depends on the magnitude *and direction* of the applied electric field, the phase change in the left guide is the same as that in the right, but of opposite sign. The application of a single voltage thus gives a "push-pull" effect, i.e., the phase is advanced in one guide and retarded in the other.

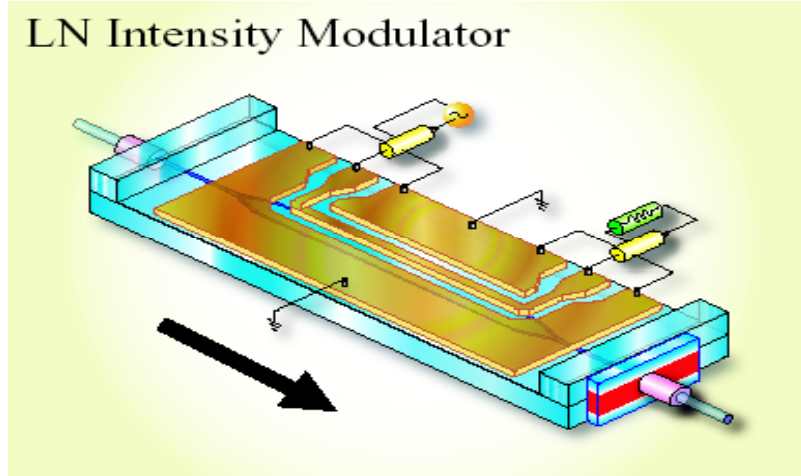
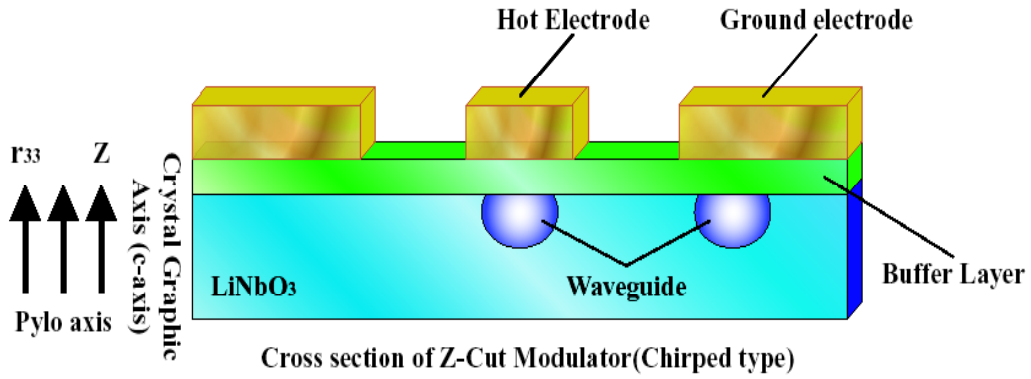


Figure 12. A single-sided z-cut Mach-Zehnder modulator.

The salient point here is that the SFDR of a link using a single-sided drive is 4 dB *worse* than one using a push-pull drive. This result, which is not intuitively obvious, was confirmed both experimentally and analytically. The poorer performance is related to the fact that the single-sided device is “wasting” the light that is not phase-shifted; the push-pull drive produces 6 dB more signal power than the single-sided drive for a given input optical power, which translates into a 4 dB advantage in the SFDR. Interestingly enough, the SFDR of an IMDD link does not depend upon the type of modulator used; the effect only occurs in coherent links.

A similar effect occurs when one uses a single detector instead of a dual-balanced device. In this case, however, the SFDR penalty is only 2 dB. This is because a 6 dB signal reduction is now accompanied by a 3 dB reduction in the noise, giving a 3 dB drop in the SNR, and a 2 dB drop ($3 \times 2/3$) in the SFDR. These results, together with a more obvious change that occurs with a 5 dB reduction in laser power, are summarized in Table 2.

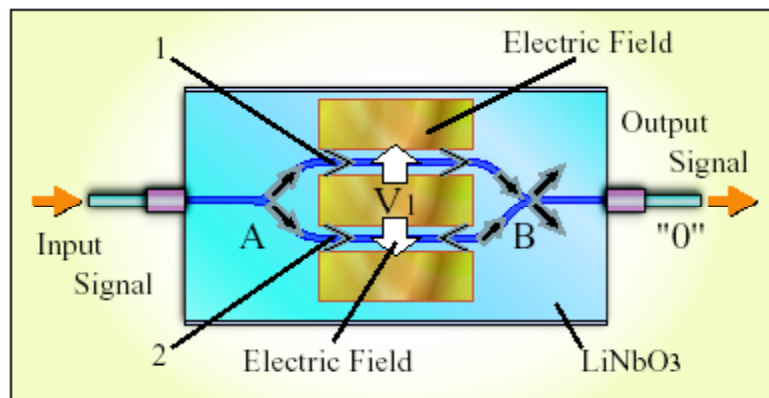
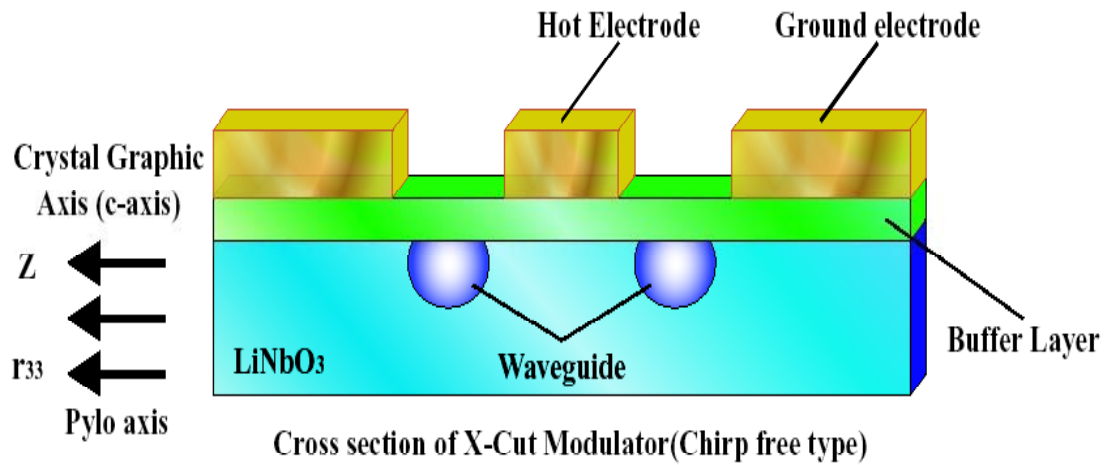


Figure 13. A push-pull (x-cut) Mach-Zehnder modulator.

Table 2. SFDR configurational handicaps.

Equipment	SFDR Handicap (dB)
Single-sided modulator (vs dual)	4.0
Single detector (vs dual-balanced)	2.0
100 mW laser (vs 300 mW)	3.2

3.4 SFDR Measurement Data

The SFDR measurements were made over the entire length of the program, not because the measurement took that long, but because improvements in components and techniques were constantly being made as time progressed. The initial measurements, for example, were made with a single-sided modulator and a single photodetector, whereas subsequent measurements were made with the more appropriate devices.

There were some setbacks along the way. The dual-balanced detector we ordered early in the program was destroyed by connecting it to too high a voltage, and had to be replaced. The replacement unfortunately had a built-in amplifier with a low IP3 that seriously degraded the SFDR of the link, and had to be swapped for one without an internal amplifier. Several problems also occurred with our 100 mW laser, which had an intermittently-defective thermal controller that allowed the laser to drift to other frequency modes; this hampered locking and made stable measurements difficult. The device had to be sent back to the manufacturer twice before the defective controller was finally repaired. However, once these issues were resolved, the experiment gave accurate and reproducible results.

Problems with thermal and environmental instabilities that we had encountered on an earlier program did not occur with this setup. We attribute this dramatic improvement in stability to the use of fusion splices and PM fiber and components throughout the link. It is our belief that the problems encountered earlier were due to small reflections at the connectors and polarization wander due to time-varying stresses in the fibers. The use of fusion splices and PM eliminated both of these problems.

Two different configurations were measured. The first, shown in Figure 6, consisted of a 100 mW laser and all of the attendant components described earlier. The second was the same link with the same laser followed by an optical amplifier (see Figure 16). The optical power available from this MOPA (Master Oscillator Power Amplifier) combination was 1 Watt. However, the highest power used during this program was 500 mW.

Measurements were made by first bringing the system up slowly, with constant readjustments of the modulator bias point so that good pinch-off (extinction) was maintained. This is a slightly hair-raising procedure, because if an operator-error occurs, one can easily zap the detector with a Watt of optical power. However, once the operating point is reached and the system stabilizes, the bias point holds rather well. For a stand-alone system, however, one would have to develop some control circuitry to do this start-up, and to lock the modulator bias point to the desired value.

The system was typically allowed to stabilize for an hour or so. The fundamental tones and 3rd-order intermodulation tones were then measured as a function of signal source power. The measurement of the fundamental tones was almost instantaneous, but measurement of the intermods was not because of proximity to the noise floor. A series of measurements was made at each frequency, each at the lowest possible resolution bandwidth, and with signal averaging used. When the link incorporated a postamplifier, a YIG filter was used to first measure the fundamental tones, and to then block these tones before they could enter the postamplifier and

spectrum analyzer, thus preventing undesired intermodulation in these devices. This required a manual tuning of the filter at each frequency. The final averaged data were plotted, fit with straight lines on a log plot, and the SFDR calculated from the intercept of these lines with the measured noise floor.

Figure 14 shows the results of one of these measurements for a laser power of 100 mW. The power of the primary tone, the intermodulation tone, and noise are plotted as a function of the source power, all in dBm. Although the laser power was 100 mW, only 26.7 mW would actually reach the detectors if the modulator were biased for full transmission (which, of course, it never is). This value, when converted to a photocurrent, determines the SFDR of the system. In addition to the components shown in Figure 6, the setup included a postamplifier with a noise figure of 3.8 dB, a gain of 45 dB, and a 1- dB compression point of 34 dBm . The noise floor in Figure 14 has been corrected for the measured kT noise of the amplifier and spectrum analyzer (a 0.6 dB correction). The total local oscillator current (I_{lo}) was 15 mA.

The measured SFDR for this case is $119.3 \text{ dB Hz}^{2/3}$. The theoretical value (equation (14)) for the measured I_s of 20 mA is $120.1 \text{ dB Hz}^{2/3}$. Our measured value is thus about 0.8 dB shy of this prediction.

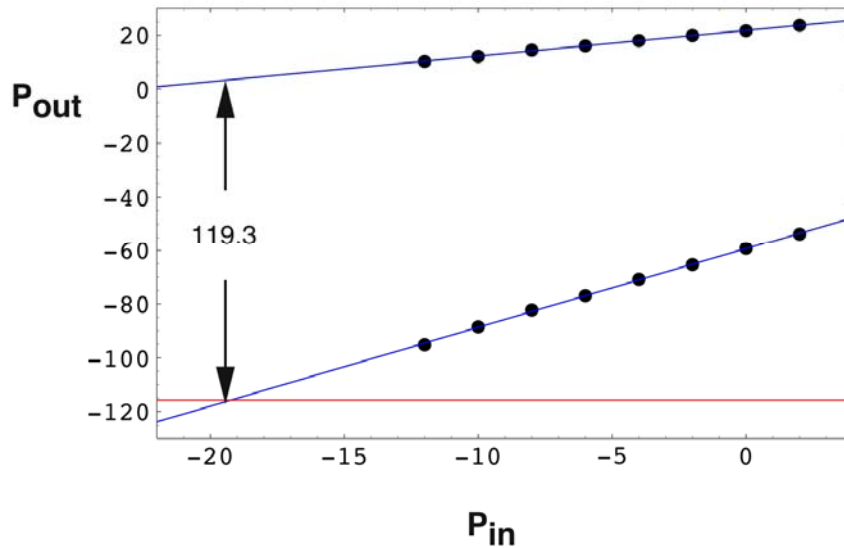


Figure 14. The fundamental (upper) and IM3 (lower) tones for a laser power of 100 mW and a total local oscillator photodetector current of 15 mA. The red line is the measured noise in a 1 Hz bandwidth.

The local oscillator current was then reduced to 5 mA by adjusting the variable splitter shown in Figure 6. The laser power remained fixed at 100 mW. The measured SFDR, shown in Figure 15, is $120.0 \text{ dB Hz}^{2/3}$, a value 0.7 dB higher than the measurement of Figure 14. Some of this increase is due to the fact that less power is going to the local oscillator, and more to the

primary signal. However, the measured value is still a bit less than the theoretical prediction of $120.55 \text{ dB Hz}^{2/3}$ that one would expect for an I_s of 24.2 mA.

$120 \text{ dB Hz}^{2/3}$ is a very respectable value, but 125 would be better. At this point we decided to try using an optical amplifier to boost the laser power to higher values. A Keopsys 1-Watt Erbium Doped Fiber Amplifier (EDFA) was available for short-term loan. This device had not been optimized for our particular low-gain application, so there was no guarantee that the noise performance would be optimum. However, some preliminary calculations showed that using 100 mW of input power should give an amplifier ideality factor of roughly two, and perhaps a little less, for the output powers we planned to use. This should provide adequate noise performance.

The experimental configuration was modified as shown in Figure 16. The amplifier was placed before the modulator, but after the first power divider. This ensured that the local oscillator signal, which undergoes only minor attenuation, was not corrupted by the EDFA noise. The EDFA gain was raised slowly from 1 to approximately 5, at which point the measured power into the modulator was 500 mW (!). The SFDR data for this configuration is shown in Figure 17.

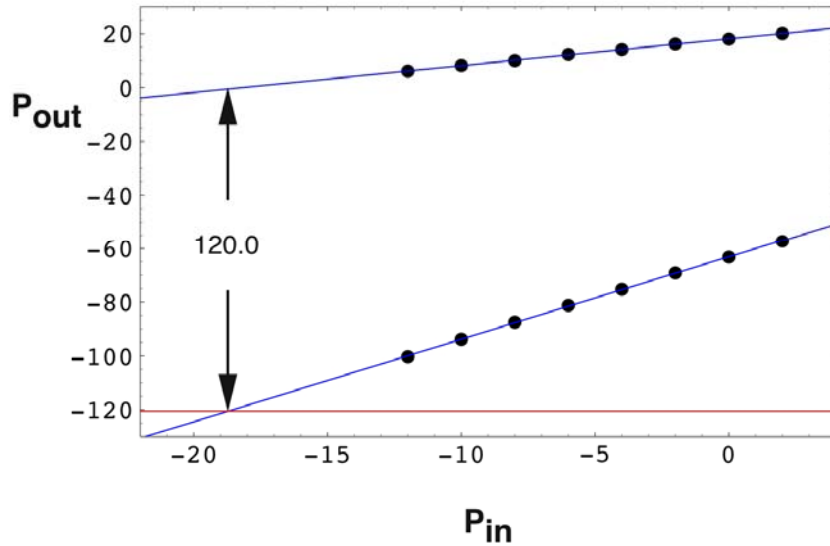


Figure 15. Measured tones and noise for a laser power of 100 mW and total local oscillator photodetector current of 5 mA.

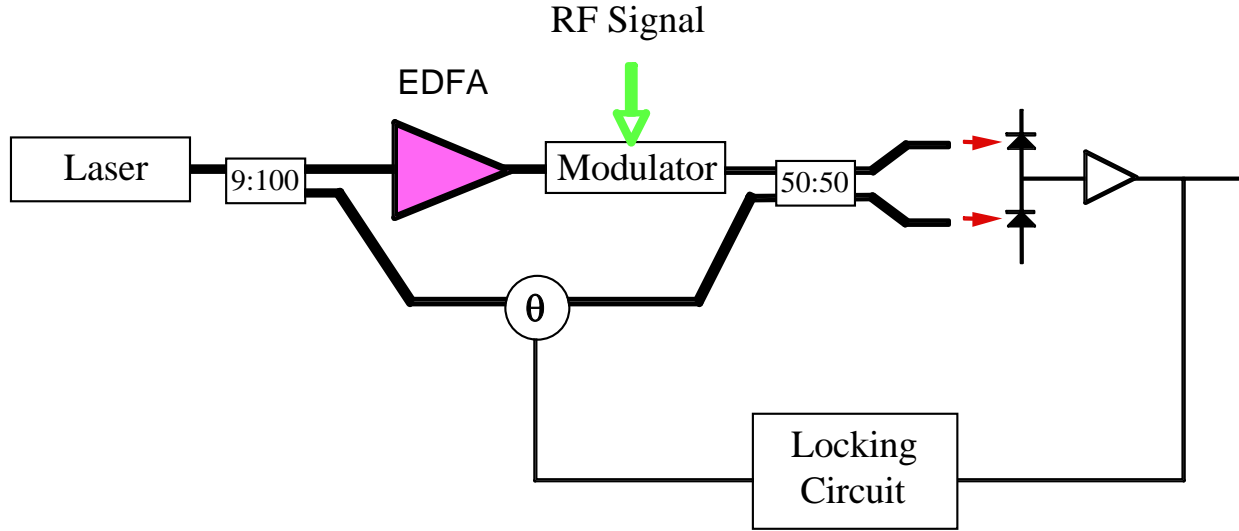


Figure 16. Link with an optical amplifier (EDFA) to boost the laser power.

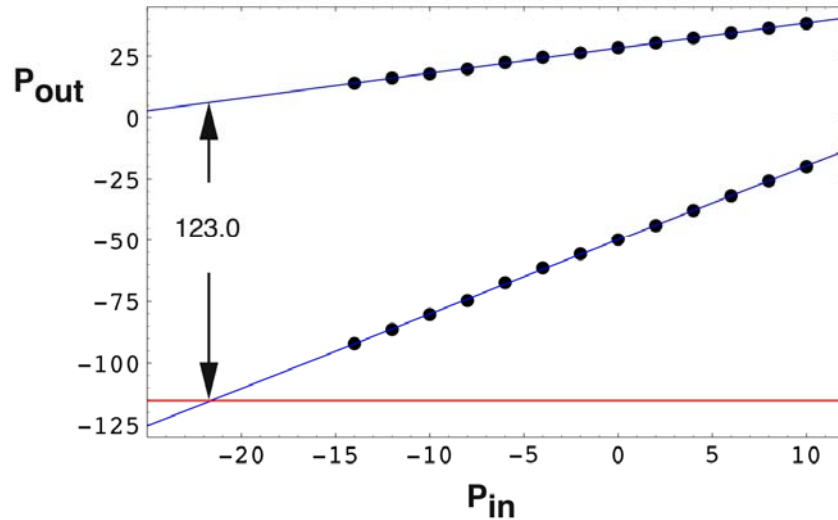


Figure 17. Measured tones and noise for the link with 500 mW of MOPA power and a total local oscillator photodetector current of 15 mA.

The measured value of 123.0 dB Hz^{2/3} is an impressive result, but is 2 dB lower than the theoretical prediction of 125.0 that one should get for an I_s of 112 mA (the measured value). We thought at first that this discrepancy might be due to a higher-than-expected amplifier noise figure, but measurements showed that the noise floor was exactly what one would expect for a 15 mA local oscillator current. Because the noise is what it should be, the cause of the lower-than-expected SFDR must be due to an unanticipated nonlinearity elsewhere in the link. It is our current belief that this additional nonlinearity is caused by the photodetectors.

3.5 Photodiode Nonlinearity

The nonlinearity of p-i-n photodetectors has been measured and analyzed by several groups. The distortion that occurs at higher photodetector currents can be separated into two types [1, 2]: (1) a component related to transit-time effects that increases linearly with increasing modulation frequency (which shows up primarily in the 2nd harmonic), and (2) a component due to the intrinsic nonlinearity of the DC responsivity that has the same frequency dependence as the detector itself, (constant at lower frequencies with a roll-off in response at the band-limiting frequency of the device).

The distortion that we are interested in can consist of both types. To keep the analysis simple, however, we shall assume here that our detectors can be described by the second type. This allows a power series analysis that leads to closed-form result.

The DC responsivity of a photodiode can be expressed as a power series

$$I = aP + bP^2 + cP^3 + \dots \quad (17)$$

where P is the applied optical power and I is the generated photocurrent. The coefficients a , b and c are normally determined by experimental measurement.

In order to determine the various dependencies, one substitutes an expression for the total optical field seen by the photodetector

$$P = \left(E_s \cos(\omega_c t) \theta(t) + E_{lo} \cos(\omega_c t + \phi) \right)^2 \quad (18)$$

per equations (1) and (3), and uses the two-tone modulation relation

$$\theta(t) = z \left(\cos(\omega_1 t) + \cos(\omega_2 t) \right) \quad (19)$$

where z is the fractional modulation amplitude.

After expanding this through third order in P , reducing the trigonometric functions to multiple-angle terms, and substituting $E_s = \sqrt{2I_s}$ and $E_{lo} = \sqrt{2I_{lo}}$, one finds that the 3rd order intermodulation term having a cubic dependence on the modulation amplitude, z , is given by

$$I_{IM3} = \left(\frac{9}{2} b \sqrt{I_{lo} I_s^3} + \frac{75}{2} c \sqrt{I_{lo}^3 I_s^3} \right) z^3 \cos(2\omega_1 t - \omega_2 t) \quad (20)$$

This is an interesting result for several reasons. First, it says that one can have 3rd-order intermodulation when the diode nonlinearity is only quadratic ($c = 0$). This is quite different from the IMDD case, which requires a finite c for I_{IM3} to exist. Second, the existence of two terms with different dependencies means that these terms can interfere and possibly cancel if the signs of b and c are opposite. A measurement of the intermod as a function of I_{lo} would then

have a sharp null for some particular value of I_{lo} . We have looked for this effect but have not found it, which implies either that b and c have the same sign, or that one term dominates.

If the b -term dominates then

$$P_{IM3} \sim I_{lo} I_s^3 \quad (21)$$

If the c -term dominates

$$P_{IM3} \sim I_{lo}^3 I_s^3 \quad (22)$$

To determine which of these terms, if either, was playing a role, we varied I_{lo} and I_s independently. If detector nonlinearity were not playing a role and the intermod were being generated by the modulator only, then the intermod power should have varied linearly with either current, i.e.,

$$P_{IM3} \sim I_{lo} I_s \quad (23)$$

We looked for the c -term first by varying I_{lo} while holding I_s constant; the intermod power changed linearly with I_{lo} , which implies that the c -term is not significant. We then varied I_s while holding I_{lo} constant, and saw an effect. The power of one intermod increased by 1.93 dB for every 1 dB change in I_s (the signal power), the power of the other by 1.33 dB per dB. Although the fact that one changed more than the other is a bit puzzling, the fact that the slope is not 1 dB per dB is significant. That it is not 3 dB per dB as predicted by equation (21) could mean that the change is a combination of the linear effect of equation (23) and the cubic effect of equation (21).

If the b -term is indeed causing the SFDR to saturate at some fixed value, then we should be able to analytically determine that value in the limit of large I_s and I_{lo} . The analysis for our dual-balanced detector is straight-forward, and gives the following relation:

$$SFDR_{diode} = \frac{0.924}{(e \Delta f b)^{2/3}} \quad (24)$$

Note that the SFDR is independent of either I_s or I_{lo} . This peculiar result is due to the fact that we have ignored thermal noise by assuming that the shot noise, $2eI_{lo}$, is much larger than kT/Z_o , a perfectly valid assumption for the high detector currents that we are considering.

It is instructional to evaluate this expression using the data for the photodiode of reference [1]. One finds that the low-frequency value of b for the particular photodiode measured in that reference is $0.448 \text{ Ampere}^{-1}$. Using this to evaluate equation (24) yields

$$SFDR_{diode} = 127.3 \text{ dB} \quad (25)$$

Thus, if our coherent AM link used the particular photodiode measured in reference [1] in a dual-balanced configuration, the highest SFDR that we could *ever* attain, assuming unlimited laser power, would be 127.3 dB Hz^{2/3}. Although our analysis is somewhat simplistic in that it ignores the high-frequency transit-time effects that occur for the 2nd order term (which could reduce this value to 124 dB!), the conclusion is inescapable: it is highly likely that there is a problem with detector nonlinearity at higher signal currents (I_s), even though the average magnitude of the signal current at the detector is small.

This result, achieved at the very end of the program, changes the complexion of the problem we were addressing. The original assumption was that the modulator was the component limiting the values of SFDR that were achievable, and that the nonlinearities of other devices were negligible. And this was true. However, the work-around provided by the coherent approach, which greatly mitigated the nonlinearity of the modulator, brought us to the point where the “negligible” distortion of these other components is now non-negligible. We shall see another example of this in the next section on linearized modulators.

4.0 Linearized AM Modulators

4.1 Introduction.

There are a variety of ways to linearize optical modulators. Some of these use a secondary modulator to generate a distortion term that is subtracted from that of the primary modulator, thereby removing that order of distortion. Others exploit a particular device feature and achieve linearization by a simple change in an operating point. The third class of devices are those that can be linearized with an approach that mimics Fourier series decomposition. And finally, there is a fourth class of devices that can be linearized via mathematical modeling, but that are very difficult to understand intuitively. A disproportionate number of linearization schemes fall into this last category.

In our proposal we discussed in some detail three particular approaches:

1. the Dual Parallel Mach Zehnder Modulator (DPMZM),
2. the Cascade Modulator, consisting of a series cascade of two Mach-Zehnder-like phase shift sections separated by directional coupler sections, and terminated in a Y combiner (DMZDCY), and
3. the Dual-Y Directional Coupler Modulator, a Y-fed directional coupler modulator terminated in a Y combiner (DYDCM).

During subsequent negotiations it was decided that only one of these three approaches would be pursued in this program: the Dual-Y Directional Coupler Modulator. The advantages of this approach were that it resulted in a small, self-contained device needing no external hardware and requiring no adjustments. A disadvantage was the consequence of this last advantage: because there were no adjustments that one could make, the device had to be fabricated with very precise tolerances in order to provide the desired functionality. Another disadvantage was that the device would have a switching voltage (V_π) that was roughly 3 times larger than comparable unlinearized devices (a disadvantage, incidentally, that is shared by many linearization techniques).

Because of this second disadvantage, we analyzed yet a fourth approach mid-way through the program: the Synthesized Directional Coupler, a device that had been studied by a group at the University of Minnesota [3]. We originally thought that this approach could yield a device that did not have a large switching voltage. In this, unfortunately, we were wrong. The approach did yield some improvement in switching voltage, but not the factor of three that we had originally anticipated. Because of this, and because the fabrication of the synthesized device would be somewhat more complex than the Dual-Y Directional Coupler Modulator, we decided to stay with the DYDCM. However, the synthesized approach had many interesting features, one of which is a much greater flexibility in the type of device that one could make, and is thus of considerable interest on its own. In the following section we shall discuss the analysis that we

did on this device as part of this program. This will be followed by a complete analysis of the DYDCM, and the experimental work we did to verify its performance.

4.2 The Synthesized Directional Coupler Modulator

4.2.1 Theory

The synthesis approach developed by Gopinath's group at the University of Minnesota allows one to, in principle, design a modulator that can have an arbitrarily-complex transfer curve [4]. The approach resembles Fourier synthesis, in that additional lengths of directional coupler are used to provide the higher Fourier spatial frequencies needed to synthesize a particular transfer curve. Thus, the more complex the transfer curve, the longer the modulator.

A directional coupler consists of two optical waveguides that are sufficiently close to each other that a small portion of the optical field in one waveguide extends into and overlaps with the mode field of the other, and visa versa. This modal overlap allows energy transfer between the two waveguides. Thus, if one launches light into the end of one guide only, this light will slowly transfer into the other guide as it travels along the coupler length. If the coupler is fabricated correctly, all of the light can leave one guide, and end up in the other. If the coupler is long enough, this transfer process will repeat, and all of the energy will transfer back to the original guide.

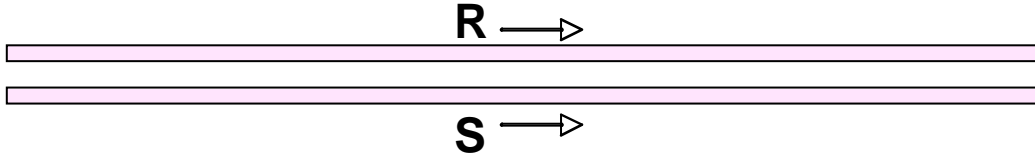


Figure 18. Two closely-space optical waveguides that form a directional coupler.

The coupled mode equations that govern this energy transfer are given by

$$\frac{dR}{dz} = j\delta R - j\kappa S \quad (26)$$

$$\frac{dS}{dz} = -j\delta S - j\kappa R \quad (27)$$

where 2δ is the propagation constant difference between the two waveguides that is caused by some external stimulus, (e.g., an electric field for an electrooptic material), κ is the coupling constant between the waveguides, and R and S are the z -dependent complex numbers that characterize the field amplitude in the waveguides. The spatial dependence of the field amplitude is given by

$$\Psi(x, y, z) = R(z)\varphi(x, y) \quad (28)$$

where $\varphi(x, y)$ is the optical mode profile. Determining R and S along the z -axis determines the fields everywhere.

One can simplify these equations by substituting

$$R \rightarrow R e^{j\delta z} \quad S \rightarrow S e^{-j\delta z} \quad (29)$$

which yields

$$\frac{dR}{dz} = -j\kappa S e^{-2j\delta z} \quad (30)$$

$$\frac{dS}{dz} = -j\kappa R e^{+2j\delta z} \quad (31)$$

To go forward with this approach, one now allows the coupling constant to have variation along its length (we'll say more about how this is done later). In the weak coupling approximation R is approximately unity and S is almost zero. When this is the case, one can write

$$\frac{dS}{dz} = -j\kappa e^{+2j\delta z} \quad (32)$$

The solution to this is

$$S(\delta) = -j \int_{-L/2}^{L/2} dz \kappa(z) e^{+2j\delta z} \quad (33)$$

In the limit of large L , where L is the length of the device, or for couplers for which $\kappa \rightarrow 0$ beyond a certain length, one can write this as

$$S(\delta) = -j \int_{-\infty}^{\infty} dz \kappa(z) e^{+2j\delta z} \quad (34)$$

Thus, S is the Fourier transform of κ .

Physically, $S(\delta)$ is the optical field amplitude at the output of the bottom guide of Figure 18 when all of the incident light is put in the top guide. Because $S(\delta)$ is the Fourier transform of $\kappa(z)$, $\kappa(z)$ will be the inverse Fourier transform of $S(\delta)$.

$$\kappa(z) = \frac{-j}{2\pi} \int_{-\infty}^{\infty} d\delta S(\delta) e^{-2j\delta z} \quad (35)$$

We now have all the tools needed to synthesize a modulator having a transfer curve shape of our choice. The general idea is as follows. For an electrooptic material such as lithium niobate, δ will be proportional to the voltage used to drive the modulator.

$$\delta = \gamma V \quad (36)$$

Thus one can write

$$S(V) = -j \int_{-\infty}^{\infty} dz \kappa(z) e^{+2j\gamma V z} \quad (37)$$

If one chooses $\kappa(z)$ correctly, one can make the amplitude of S depend on V however one wants. For example, suppose one would like to have the triangular response shown in Figure 19. What must the functional form of $\kappa(z)$ be to get this shape?

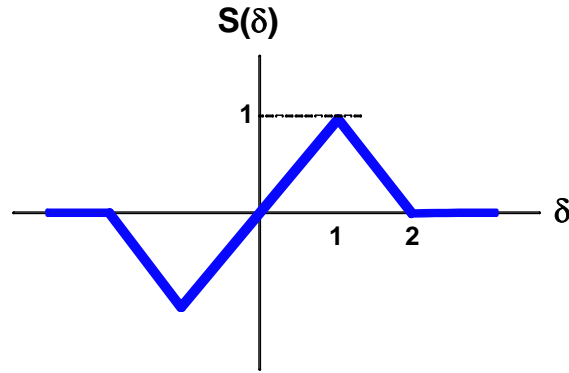


Figure 19. A triangular modulator transfer curve.

Using equation (35) one finds that $\kappa(z)$ must be

$$\begin{aligned} \kappa(z) &= \frac{2}{\pi} \int_0^{\infty} S(\delta) \sin(2\delta z) d\delta = \frac{2}{\pi} \int_0^1 \delta \sin(2\delta z) d\delta + \frac{2}{\pi} \int_1^2 (2 - \delta) \sin(2\delta z) d\delta \\ &= \frac{4 \cos z \sin^3 z}{\pi z^2} \end{aligned} \quad (38)$$

This is plotted in Figure 20 below. We shall postpone the issue of how one gets a negative κ , and assume for the moment that it can be done. The question for the moment is: how well would a modulator, made with such a κ profile, reproduce the desired shape shown in Figure 19? To find out, one cannot simply take the inverse transform of equation (34). This whole procedure is, at best, an approximation. The correct way to determine this is to use this κ profile, and find the exact S that would result by solving the original differential equations ((26) and (27)) exactly (numerically).

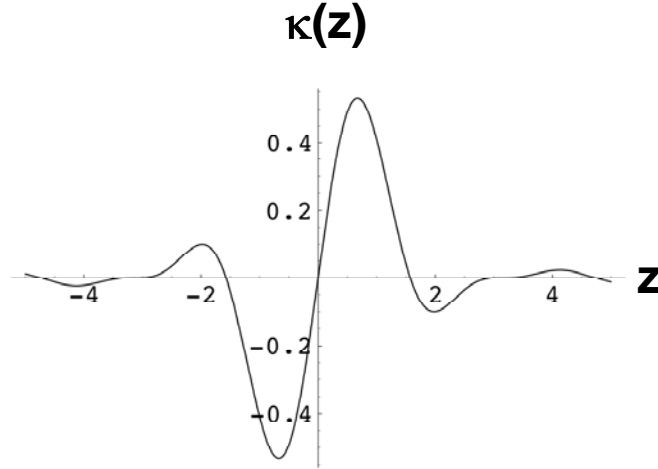


Figure 20. The shape of $\kappa(z)$ needed to get the profile of Figure 19.

This has been done using Mathematica's NDSolve routine and the initial conditions $R_o = 1$ and $S_o = 0$ (i.e., all of the input light in the top guide, none in the bottom). The result, shown in Figure 21, looks remarkably triangular, given the approximate nature of the derivation.

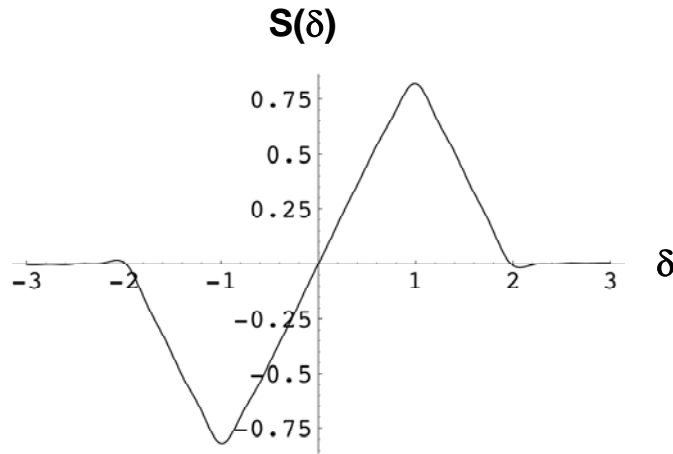


Figure 21. The transfer curve that results from the κ of Figure 20.

Unfortunately, the device length needed to get this shape and still have a reasonable switching voltage is roughly 20 cm, which is quite a bit larger than the 10 cm wafer size that is commercially available. However, a high-linearity modulator need not have a perfectly triangular shape; it should, however, have very high linearity near the origin.

After some experimentation it was found that one could increase the slope efficiency by increasing κ , and could eliminate the 3rd order term by carefully adjusting the length of the coupler. For a device length of 7.83 cm and a larger κ , one gets the response shown in Figure 22.

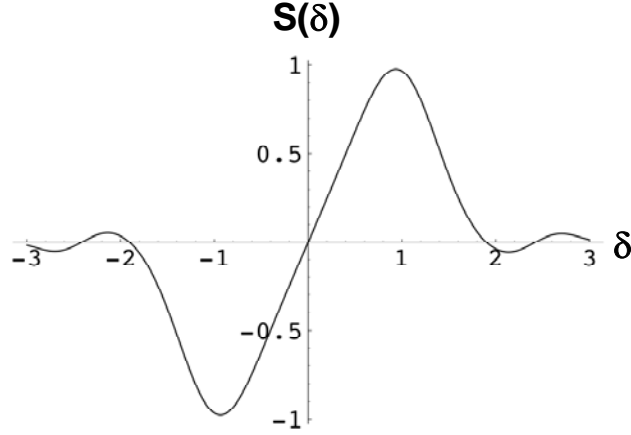


Figure 22. A linearized device with the 3rd order term eliminated and a 7.83 cm length.

We define the Slope Efficiency to be the change in transmissivity with voltage at $V = 0$.

$$\text{Slope Efficiency} = \frac{dS}{dV} \quad (39)$$

This is a measure of the change in optical throughput for a given change in voltage. The larger the slope efficiency, the more efficient the device. To get the slope efficiency, we used our measured values of $V_\pi L$ for our Mach-Zehnder modulators, and the relation

$$\delta = \frac{\pi V}{2V_\pi L} \quad (40)$$

For a $V_\pi L$ of 16 Volt-cm, the slope efficiency for this particular example is 0.12/Volt.

This synthesized design, although of some academic interest, is of no immediate practical use for two reasons: (1) κ must vary along the coupler length in a well-controlled fashion, which is hard to do accurately in LiNbO_3 and (2) κ must change sign, which is impossible.

The Minnesota group found an ingenious solution to this second problem: they showed [3] that a sign change in κ can be affected by shifting the phase in either guide abruptly by 180° . We have checked this analytically, and found that this is indeed the case. Achieving an accurate variation of κ along the coupler length, however, is harder to do. One can envision schemes in which one could ion-mill a slot of varying depth between the waveguides, or fabricate the waveguide with a variable separation distance. These approaches, however, would be difficult to implement because of the sensitivity of both to the actual titanium diffusion process that forms the waveguides.

A more practical approach is to use a constant κ , and to approximate the desired $\kappa(z)$ profile with a series of rectangular steps, each with the approximate area of the synthesized $\kappa(z)$, and with 180° phase-changes inserted wherever one wanted one rectangle to end and the next to begin. This is shown schematically in Figure 23.

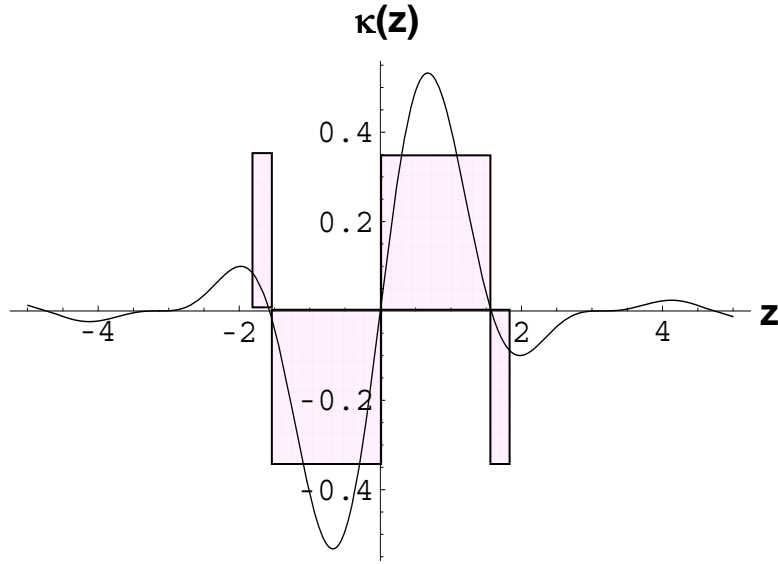


Figure 23. A device design that uses a constant κ and abrupt phase changes to achieve linearization.

The device has three phase changes: one between each rectangle and the next rectangle. The device ends where the rectangles end. The value of κ , constant along the coupler length, was chosen so that the area approximated that of the $\kappa(z)$ curve that it overlays. The width of the shorter rectangles at the ends was carefully adjusted to eliminate the third order term in the resultant transfer curve. Thus, the larger rectangles provide the general shape of the transfer curve, the smaller “corrector” rectangles provide the linearization.

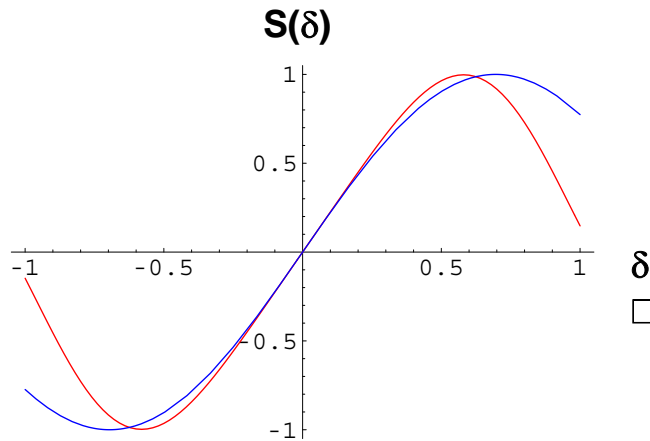


Figure 24. Transfer curve for the rectangular κ -profile of Figure 23 (red) and for a Mach-Zehnder modulator with the same slope efficiency (blue).

Figure 24 shows the transfer curve of the results for the rectangular profile of Figure 23 when the corrector sections are adjusted to eliminate the 3rd order term. Note that the visible

improvement over that of an unlinearized device is not that impressive. However, the change in SFDR, as we shall see shortly, is dramatic.

The primary reason for investigating this synthesis approach was to see if we could get a high degree of linearization while still maintaining a reasonable slope efficiency. To do this, we developed a design approach that allowed us to both maximize the slope efficiency and eliminate the 3rd order term for a given device length. The procedure was as follows:

1. Choose the length of the primary section, L (i.e., the combined length of the two larger rectangles).
2. For a given κ , adjust the “corrector” lengths so that the 3rd derivative of the transfer curve was zero. This eliminates the third-order term of the transfer curve.
3. Repeat this process for incremental steps in κ , and find the value that gives the maximum slope efficiency.

A plot of the slope efficiency that one gets when this process is carried through is shown in Figure 25.

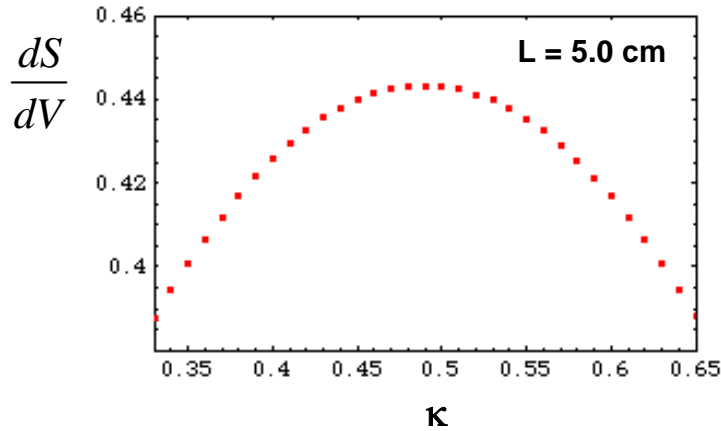


Figure 25. Slope efficiency versus κ . Each point on the curve is for a device design that has been linearized. The optimum value is 0.48.

A good figure of merit for these devices is the slope efficiency per unit modulator length, defined by

$$\frac{SE}{L} = \frac{1}{L} \frac{dS}{dV} \quad (41)$$

Our analyses showed that there are two different values for each modulator design: one for the unlinearized device, and one for the linearized. For our rectangular constant- κ approach, one finds that linearizing reduces the slope efficiency by roughly a factor of two. The values are shown in Table 3, together with that for an unlinearized Mach-Zehnder modulator.

Table 3. Slope efficiencies for various configurations.

TYPE	SE/L
Unlinearized	0.071 /Vcm
Linearized	0.037/Vcm
MZ modulator	0.098/Vcm

To formulate a design that would give the best performance consistent with our existing manufacturing capabilities, we choose a device length that would fit on a 3" (7.6 cm) diameter LiNbO3 wafer. The final parameters were:

- a primary section length of 5 cm
- a total device length of 6.036 cm
- an optimum κ of 0.48
- a slope efficiency of 0.22/Volt

4.2.2 SFDR

The SFDR for this design was calculated by first expanding S in a power series, realizing that the 3rd order term will be missing because of the linearization procedure.

$$S(\delta) = a_1 \delta - a_5 \delta^5 + \dots \quad (42)$$

Using this, and the definition of the SFDR, one finds that the two-tone SFDR for any modulator with no 3rd order term is given by

$$SFDR = \left(\frac{I_s}{e\Delta f} \right)^{4/5} \left(\frac{8}{25 a_5} \right)^{2/5} a_1^2 \quad (43)$$

For the design outlined above, a_1 and a_5 are 2.257 and 7.03, respectively. Using these values and an I_s of 100 mA, one finds that

$$SFDR = 144 \text{ dB Hz}^{4/5} \quad (44)$$

This is substantially better than the 124.6 dB Hz available with an unlinearized Mach-Zehnder modulator.

4.2.3 Manufacturing Tolerances

The 3rd order term of the transfer curve was eliminated by carefully adjusting the corrector lengths until it disappeared. We have evaluated the error in the corrector length and the error in the three 180° phase reversers that can be tolerated before the device loses functionality. The criterion used was that the modulator should provide at least 7 dB of SFDR improvement over an unlinearized system. Table 4 shows the results of this analysis.

Table 4. Manufacturing tolerances.

Parameter	Tolerance
Corrector Length,	0.0081 cm
Phase reversers (IM3)	11° (all three)
Phase Reversers (2 ω)	TBD
κ	very insensitive

4.2.4 Phase Reversal Sections and Fabrication Approach

Achieving an almost instantaneous phase change of 180° is not trivial. One approach with a reasonably good chance of success would be to form a trench in one waveguide using ion-milling, as shown in Figure 26. The trench would then be filled with a dielectric having a slightly different index of refraction, such that

$$\Delta n x = \frac{\lambda_o}{2} \quad (45)$$

If the optical loss proved negligible, air itself could be used, thereby eliminating one processing step.

The most critical part of the fabrication process would be attaining the correct “corrector” length for a given κ . The physical length itself is not a problem; lithographic accuracy is such that one could manufacture this with micron accuracy. The problem is that the corrector length must be a given fraction of a *coupling* length, not a physical length. This is a problem, because the coupling length is a strong function of the amount of titanium used to form the in-diffused optical waveguides, and of the exact distance between the waveguides. Fabrication of directional couplers on another program has shown that reproducibly attaining a specific coupling length to better than 5% is difficult.

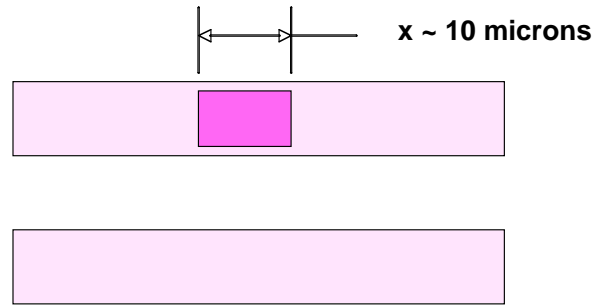


Figure 26. Details of phase-reversal trench.

One could, however, demonstrate a synthesized modulator on single devices by using a “hand-tailored” approach. This would entail fabricating long directional coupler waveguides, installing the phase reversal trenches at the appropriate places, putting metal electrodes on the waveguides (with no ports or transitions), and then adjusting the length of the corrector sections by successively and symmetrically shortening the device length with a high-speed diamond-blade saw. Between cuts one would characterize the device’s linearity by measuring the slope of the transfer curve, and then calculate the amount to be trimmed on the next cut. We have developed lensed-fiber optical coupling techniques and have the special RF probes that allow one to optically couple into and electrically connect to two very closely spaced waveguides, and have developed cutback procedures that would allow the length adjustment to be made to the required accuracy. These procedures and techniques, developed for the Y-fed directional coupler modulators and discussed in section 5.2 and 5.4, could also be used for this approach.

4.3 Dual-Y Directional Coupler Modulators

4.3.1 Introduction

The Y-fed directional coupler modulator has several noteworthy properties. Of most interest to us is the fact that it can be linearized, and in a way that supposedly minimizes the sensitivity to manufacturing tolerances [5]. As proposed, however, the device does not suppress the carrier, and hence can only be used as an intensity modulator. However, as with the previous device, we have found that a simple modification gives fully-suppressed carrier double sideband Amplitude Modulation, and does this in a way that allows a comparable degree of linearization.

The carrier is suppressed by shifting the optical phase of one output by 180° , so that this output is inverted, and then summing the inverted and non-inverted outputs with a second Y. This technique, first suggested by Desormiere [6] for a different application, subtracts the two output fields. For balanced inputs (provided by the first Y) and no applied voltage, the two fields are equal, so that their subtraction leads to zero intensity in the absence of any modulation.

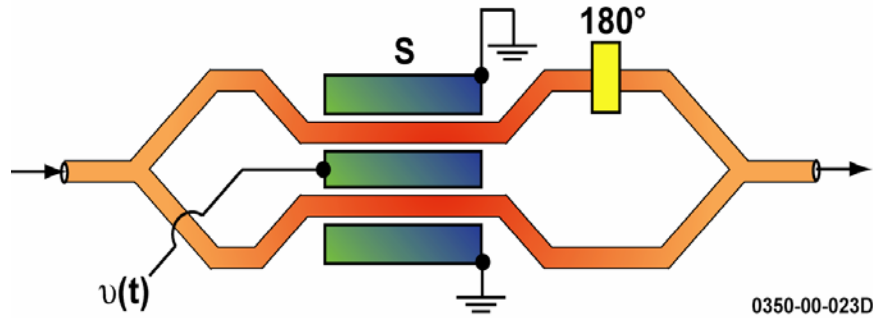


Figure 27. Dual-Y Directional Coupler Modulator.

Applying a voltage to the coupler section changes the relative indices of refraction in the two arms of the coupler. This allows quadrature components to develop in each guide that are 180° out of phase with each other. When subtracted in the final Y, the two quadrature components add, whereas the in-phase components subtract. Because the growth of the quadrature components is approximately proportional to the applied electric field, the emerging light has a amplitude that is approximately proportional to the modulation voltage. Hence, just as with the simple Mach Zehnder modulator, which cancels the carrier in a similar way, one achieves true amplitude modulation with full carrier suppression.

Linearization schemes for directional couplers are usually difficult to understand intuitively. The design procedure is often to assemble (mathematically) a series of active and passive coupler sections, and to then vary all of the device parameters until one finds a sweet spot in the transfer curve that produces some degree of linearization. The Double-Y directional coupler is different, in that one can fathom immediately what needs to be done to achieve linearization.

The optical fields at the two output ports of the directional coupler are given by the matrix equation [7]

$$\begin{bmatrix} R \\ S \end{bmatrix} = \begin{bmatrix} A & -jB \\ -jB^* & A^* \end{bmatrix} \cdot \begin{bmatrix} R_o \\ S_o \end{bmatrix} \quad (46)$$

where

$$A = \cos\left(\frac{\pi s}{2} \sqrt{1+x^2}\right) + j \frac{x}{\sqrt{1+x^2}} \sin\left(\frac{\pi s}{2} \sqrt{1+x^2}\right) \quad (47)$$

$$B = \frac{\sin\left(\frac{\pi s}{2} \sqrt{1+x^2}\right)}{\sqrt{1+x^2}} \quad (48)$$

and where x is δ/κ . δ is the change in propagation constant in each guide due to an applied field, and κ is the coupling coefficient between guides. s is the actual coupler length, L , divided by the coupling length for complete power transfer from one arm to the other, l_o .

The effect of the first Y-branch is to split the incoming optical field into two equal components. To conserve power, one must have $R_o = S_o = E_{in} / \sqrt{2}$. Inserting this in equation (2), and then subtracting the two outputs gives

$$E_{out} = \left(\frac{A - A^*}{2} \right) E_{in} = j \frac{x \sin\left(\frac{\pi s}{2} \sqrt{1+x^2}\right)}{\sqrt{1+x^2}} E_{in} \quad (49)$$

where x is proportional to the applied voltage. Hence, if $x \ll 1$, so that the quadratic terms can be ignored, and if s is an odd integer (so that the modulation is maximized), then to first order,

$$E_{out} \cong j x E_{in} \quad (50)$$

Thus the output is the modulation signal, $x(t)$, multiplied by the carrier signal, $E_s \sin(\omega_c t)$, which is the desired form for a true AM signal.

For larger values of x , the quadratic terms cannot be ignored. As x increases, the sine term decreases and the numerator increases, so that the overall value of E_{out} is lowered from the value predicted by equation (5). This lowering produces an undesired curvature in the transfer function.

Suppose, however, that the sine term were to *increase* while the denominator also increased. If the increases were equal, the effects would compensate, so that this bending away from the linear response of equation (1) would be reduced. One can in fact do just this by

adjusting s to be slightly less than the value that gives maximum slope efficiency. Increasing x now moves the sine term towards its maximum, causing it to increase. By picking the proper point on the sine curve, one can match (through a given order in x) the increase in the denominator.

To determine the optimum value of s , one expands the transfer function (49) as a power series in x . Through third order, one finds

$$E_{out} = j \left(x \sin \left[\frac{1}{2} \pi s \right] + x^3 \left(\frac{1}{4} \pi s \cos \left[\frac{1}{2} \pi s \right] - \frac{1}{2} \sin \left[\frac{1}{2} \pi s \right] \right) + x^5 \dots \right) E_{in} \quad (51)$$

As noted above, the values of s that maximize the linear term are 1, 3, 5, ... However, the values that eliminate the cubic term are found to be $s = 2.861, 4.918, 6.942, \dots$ i.e., numbers just slightly less than 3, 5, 7, etc. Thus, the shortest directional coupler that can eliminate the cubic term has $s = 2.861$. Operating at this point reduces the slope efficiency a mere 2.5 % from its $s = 3.0$ value.

Unfortunately, the value of s that removes the cubic term does not remove the fifth or higher order term. Conversely, s can be adjusted to remove the fifth order term, but only at the expense of having the cubic term return at full strength. There are schemes that will remove multiple terms, but each higher term removed brings with it an increase in device complexity and an increase in the required manufacturing precision. For reasons discussed earlier, removing the cubic term should be sufficient for this program.

Eliminating the cubic term does not have a dramatic effect on the visual shape of the transfer curve (Figure 28). It does, however, significantly increase the SFDR, especially for smaller bandwidths.

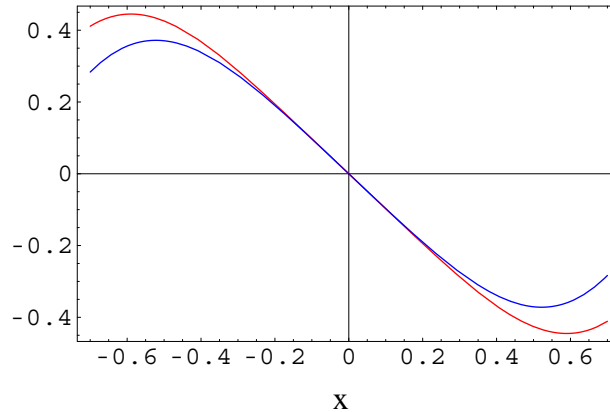


Figure 28. Transfer curve for $s = 3.0$ (blue) and $s = 2.861$ (red).

4.3.2 Sensitivity Analysis

The dual-Y directional coupler modulator has four sensitivities that must be addressed if the device is to perform as advertised. The first of these is an extreme sensitivity to the coupler length parameter, s . In the measurement section that follows, the data will show that s must be set with an accuracy of four significant figures. We shall see that this length can be adjusted by trimming the length of the electrodes, which is considerably easier than trimming the actual length of the coupling section, but even electrode trimming is not entirely trivial to do.

The second is sensitivity to the symmetry of the two Y-branches. If these do not split and recombine the fields equally, and with equal phase delay, the even harmonics and intermodulation terms reappear. Thus, for systems requiring multi-octave performance, manufacturing techniques will have to be found to trim the split values, and to remediate any differences in phase. For sub-octave applications, however, the sensitivity to split errors is much less. Errors as large as 10% in the output branch change the SFDR by less than 1 dB. Such large split errors, however, allow a substantial fraction of the carrier to pass through the device. Because of our desire to operate in a true suppressed-carrier mode, this is to be avoided whenever possible.

The third is sensitivity to differences in the microwave and optical propagation velocities. Although all modulators show a reduction in bandwidth with velocity mismatch, the dual-Y directional coupler also suffers a pronounced reduction in linearization. Fortunately, as we shall see shortly, the effect does not appear to be too severe. A reasonable velocity match should allow one to achieve 7 dB or more of SFDR improvement over the entire operational band.

The fourth sensitivity, however, cannot be so easily waved away. This sensitivity, which is peculiar to linearized directional coupler modulators, is a reduction in linearization that occurs at higher frequencies due to the microwave loss in the traveling wave electrodes. This sensitivity, which was not discovered until well into this program, is due to the fact that the “corrector” sections of the linearized couplers must cancel the nonlinearity of the primary “modulator” sections. If the cancellation signal generated by the corrector section is smaller or larger than that generated by the primary modulator section then complete cancellation cannot occur and the linearization suffers. We do not have, at this time, a solution that would allow the DYDCM to provide high linearity over multi-octave or even over a complete octave of bandwidth at frequencies above several GHz. It can, however, provide sub-octave performance.

We shall now address all of these sensitivities in some detail with the exception of the first, which is primarily a manufacturing problem that will be discussed later in some detail in the section on measurements and remediation.

4.3.2.1 Y-Branch Imbalance

It is useful to describe the optical modes in a directional coupler in terms of the even and odd modes. This is particularly true for the Dual-Y device, which operates only on the antisymmetric mode, and completely ignores the symmetric mode. Figure 29 shows the profiles of the individual optical modes in the R and S waveguides. One can convert these to the normal modes (eigenmodes) of the system using coupled mode analysis. The even and odd modes that

result are the sum (even) and difference (odd) of the individual R and S modes. These modes have different propagation velocities, with the even mode being slightly slower than the odd.

These eigenmodes are handy for a variety of reasons. One is that the first Y-branch of the DYDCM splits the light-beam and launches a pure even mode into the coupler. If no voltage is applied to the directional coupler electrodes, this even mode will remain unchanged as it propagates down the coupler, and no odd mode will form. Because of this constancy, the even and odd modes form a nice set of basis states that can be used for coupler analysis, in which changes due to an applied electric field are treated as perturbations.

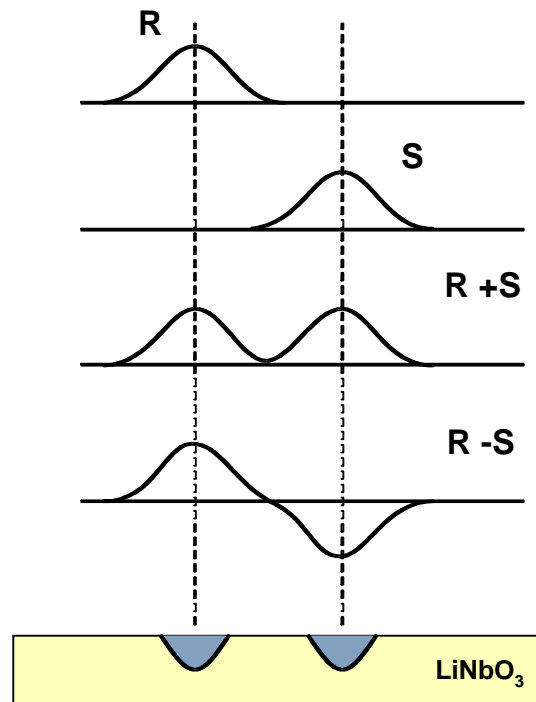


Figure 29. Even and Odd mode formalism.

Another reason is that the second Y-branch acts as a mode filter, passing the even mode and filtering out the odd mode. If one adds a 180° phase-shifter, as shown in Figure 30, the combination filters out the even mode and passes the odd (the flipped waveform will, because of its opposite sign, cancel the non-flipped waveform when the two are summed in the junction).

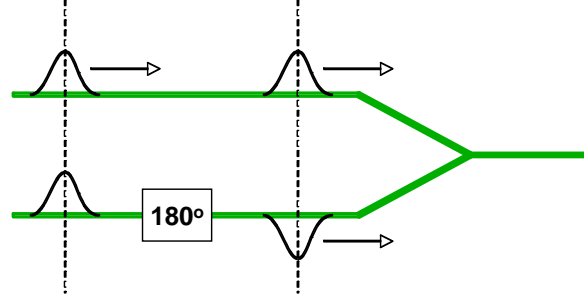


Figure 30. Mode-filtering properties of a Y-branch with a phase-shifter.

When a differential voltage is applied to the two electrodes, the velocities in the two guides are changed relative to each other. This allows the even mode and odd modes to couple. The odd mode can now rob power from the even mode, so that the odd mode grows (from zero) as it moves along the coupler, while the even mode diminishes.

If one transforms the R-S equations of the waveguides (equations (26) and (27)) to the even-odd equations, one finds

$$\frac{dE}{dz} = i\delta O \exp[+i2\kappa z] \quad (52)$$

$$\frac{dO}{dz} = i\delta E \exp[-i2\kappa z] \quad (53)$$

where E and O designate the even and odd mode, and δ and κ have the same meaning as before, and with δ being proportional to the applied voltage. Note that if the applied voltage is zero ($\delta = 0$), there is no change in either mode.

The solutions for the initial conditions $E(0) = 1$ and $O(0) = 0$ are

$$E(z) = \exp[i\kappa z] \left(\cos\left[\sqrt{\delta^2 + \kappa^2} z\right] - i\kappa \frac{\sin\left[\sqrt{\delta^2 + \kappa^2} z\right]}{\sqrt{\delta^2 + \kappa^2}} \right) \quad (54)$$

$$O(z) = i \frac{\delta}{\sqrt{\delta^2 + \kappa^2}} \exp[-i\kappa z] \sin\left[\sqrt{\delta^2 + \kappa^2} z\right] \quad (55)$$

If we evaluate equation (55) at the end of the coupler by making the substitutions

$$x = \frac{\delta}{\kappa}, \quad L_{transfer} = \frac{\pi}{2\kappa}, \quad s = \frac{z}{L_{transfer}} \quad (56)$$

and use the fact that passing through the 180° phase shifter and final Y-junction will remove the even mode but leave the odd mode unaffected, one finds that the modulator output field is given by

$$E_{out} = j \frac{x \sin\left(\frac{\pi s \sqrt{1+x^2}}{2}\right)}{\sqrt{1+x^2}} E_{in} \quad (57)$$

which is just equation (49).

The growth of the odd mode occurs only in the region where an electric field is being applied to the waveguides. If there is an additional length of directional coupler that follows the electrode section, the constancy of the odd mode in regions of no field insures that the result of equation (57) is unchanged, i.e., the additional length σ shown in Figure 31 plays no role whatsoever.

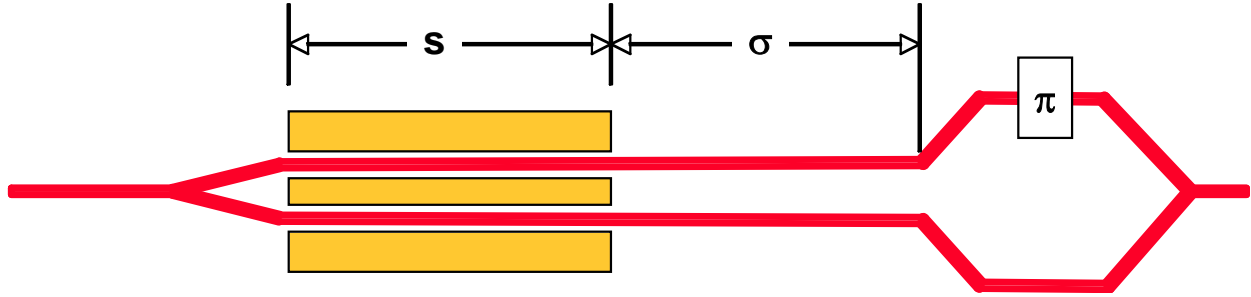


Figure 31. The electrode length, not the coupler length, determines the parameters.

This is quite different from the Y-fed directional coupler modulator used for IMDD applications. Here an additional length of coupler following the electrode section changes the output, usually for the worse. The fact that the electrode length determines s for our DYDCM is important, because it means that one can trim s using chemical etching or ablative techniques to obtain the correct value of 2.861. The IMDD Y-fed coupler requires that one physically saw off the end of the device to do such tuning.

4.3.2.2 Second Harmonic

The primary effect of phase imbalance in the input and output Y-branches is the generation of a second harmonic. Consider a Y-branch in which the fields are described as shown in Figure 32, where α and β are complex numbers that contain both phase and amplitude information, and where $\alpha \neq \beta$. If one now evaluates the output field of a DYDCM having two of these unbalanced junctions, i.e.,

$$Y_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} \alpha \\ \beta \end{bmatrix}, \quad Y_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} a \\ b \end{bmatrix}$$

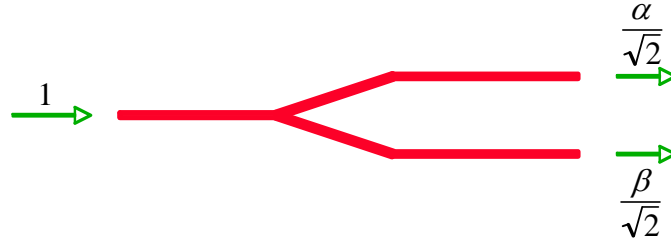


Figure 32. An unbalanced Y-junction.

one finds that the expression for the output field is

$$\frac{E_{out}}{E_{in}} = \frac{(a\alpha - b\beta)}{2} \cos\left(\frac{\pi}{2} s \sqrt{1+x^2}\right) + i \frac{(b\alpha - a\beta)}{2} \frac{\sin\left(\frac{\pi}{2} s \sqrt{1+x^2}\right)}{\sqrt{1+x^2}} + i \frac{(a\alpha + b\beta)}{2} x \frac{\sin\left(\frac{\pi}{2} s \sqrt{1+x^2}\right)}{\sqrt{1+x^2}} \quad (58)$$

If α , β , a and b were all unity, as they would be in ideal junctions, then the first two terms would disappear, and one would regain equation (57). However, if they are not, then the first two terms are non-zero, and a second harmonic will result.

There are several results that one can glean from these equations that are worth mentioning:

1. One cannot adjust the imbalance of Y_2 to fix that of Y_1
2. The first term only generates a 2nd harmonic when $\text{Im}(a\alpha - b\beta) \neq 0$ (phase imbalance)
3. The 2nd harmonic of the second term is proportional to $\text{Re}(b\alpha - a\beta)$ (amplitude imbalance)
4. The second term produces a much weaker 2nd harmonic because the quadratic dependence on x has been removed by linearization
5. The insensitivity of the 2nd harmonic to amplitude imbalance assumes perfect linearization (perhaps not realistic).

Comments 2 through 5 assume perfect homodyne detection (i.e., no local oscillator phase error).

One can calculate the SFDR due to the 2nd harmonic for terms one and two, finding

$$SFDR_{T1} = \frac{8}{\text{Im}\left(\frac{a\alpha - b\beta}{2}\right) \pi s \sin\left(\frac{\pi s}{2}\right)} \left(\frac{I_s}{eB}\right)^{1/2} \cong 89 \text{ dB Hz}^{1/2} - 10 \log_{10}[\sin \phi] \quad (59)$$

$$SFDR_{T2} = \frac{8}{\pi s \sqrt{\text{Re}\left(\frac{(b\alpha - a\beta)}{2}\right) \sin\left(\frac{\pi s}{2}\right)}} \left(\frac{I_s}{eB}\right)^{3/4} \cong 133 \text{ dB Hz}^{3/4} - 10 \log_{10}\left[\text{Re}\left(\frac{(b\alpha - a\beta)}{2}\right)\right] \quad (60)$$

In equation (59), 2ϕ is the phase difference between α and β if all the error is in Y_1 , or the average phase error per arm if it is not. The numerical expressions to the right have been evaluated for an I_s of 100 mA.

A plot of the tolerable phase error as a function of the measurement bandwidth is shown in Figure 33. The criterion for determining what is tolerable is: a ϕ that gives $SFDR_{2nd} = SFDR_{IM3}$.

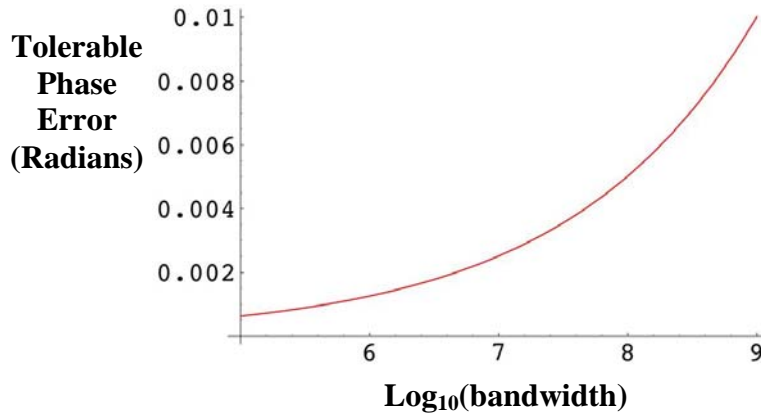


Figure 33. Input Y-branch phase sensitivity.

These results seemed rather severe. However, the Mach-Zehnder modulator has a very similar problem: one must bias the device exactly at V_π and have perfect extinction to eliminate the 2nd harmonic. In fact, the sensitivities to phase and amplitude error are numerically comparable to those for the Dual-Y Directional Coupler Modulator. All of the linearization schemes discussed here and in the proposal would require some type of fine adjustment and control to suppress the 2nd harmonic.

There are several ways in which this could be done for the DYDCM. One of these, shown in Figure 34, uses two short DC electrodes with adjustable voltages to fix both amplitude and phase errors in the Y-branches. This technique was developed on another program to trim the input of a Y-fed directional coupler intensity modulator. The only drawback to this approach is that it would require 4 connections and 4 voltages to fix both junctions.

A second scheme takes advantage of the fact that an imbalance condition can be represented by a combination of even and odd modes. One could “ungrow” the odd mode by applying a voltage at the right place along the structure. Thus, one could envision having a series of four electrodes, of which only one is used (Figure 35). One would have to determine which one experimentally, but once this was done, two voltages could be used to correct both junctions.

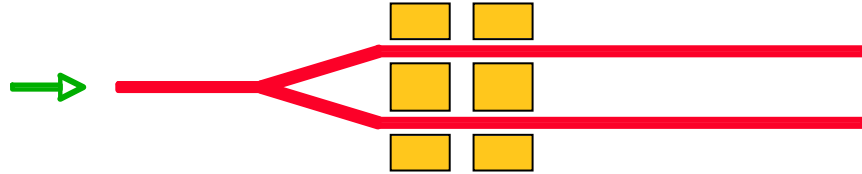


Figure 34. A dual-electrode scheme for adjusting the phase and amplitude imbalance of a Y-junction.

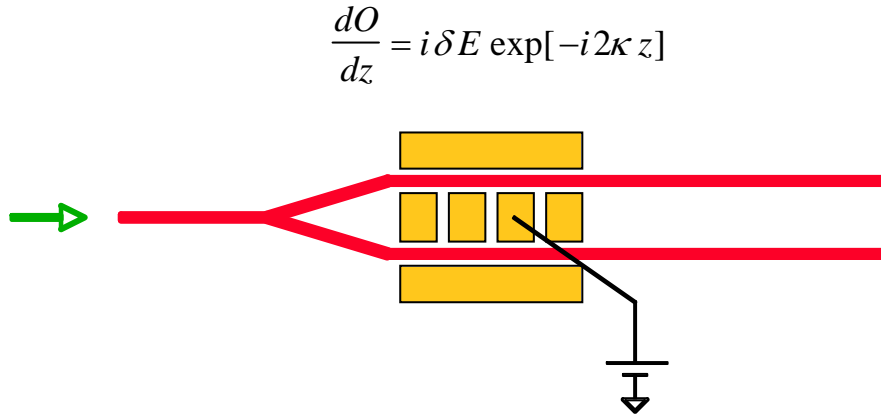


Figure 35. A second scheme for balancing the Y-junction. This uses a strategically-placed electrode (1 of 4) to “ungrow” the offensive odd mode.

The third scheme assumes that good linearization is achieved, so that the second term of equation (58) can be ignored. When this is the case, a simple phase adjustment in the phase reversal section may fix most of the unbalance. The advantage of this approach is that only one electrode is needed (Figure 36), and that electrode is the same as the one doing the 180° phase reversal.

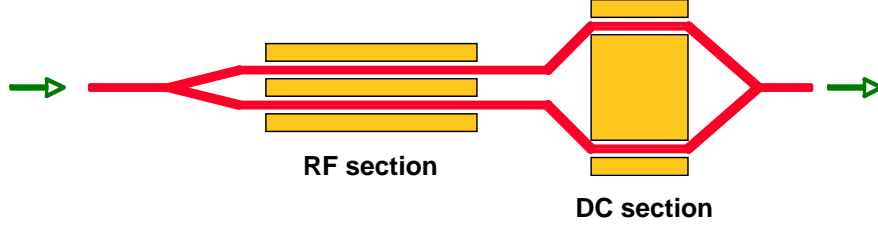


Figure 36. 2nd harmonic reduction and phase reversal using a common electrode.

4.3.3 Velocity Mismatch

There are at least two ways to quantify the sensitivity of the Dual-Y Directional Coupler Modulator to velocity mismatch. One is to use a time-domain approach described in the proposal, with propagation of the microwave and optical fields handled by step-wise discrete matrix multiplication. The second, which will be described in some detail below, is to find the coupled differential equations that govern the propagation of the optical fields in space and time, and to solve these equations by Fourier decomposition.

4.3.3.1 Traveling Wave Coupled Mode Analysis

Using a coupled mode approach such as that described in [8], one can show that the time-dependent differential equations for the optical fields propagating in each guide of the directional coupler are, in the slowly-varying envelope approximation, given by

$$\frac{\partial R}{\partial z} + \frac{n_o}{c} \frac{\partial R}{\partial t} = +i \delta \cos[\omega t - \beta z] R - i \kappa S \quad (61)$$

$$\frac{\partial S}{\partial z} + \frac{n_o}{c} \frac{\partial S}{\partial t} = -i \delta \cos[\omega t - \beta z] S - i \kappa R \quad (62)$$

R and S are the optical fields in the top and bottom guide, respectively, δ is the change in propagation constant in each guide due to an applied field, and κ is the coupling coefficient between guides. The cosine terms in each equation describe the *single-tone* traveling microwave field being applied to the two directional coupler electrodes. When β is positive, the optical and microwave fields propagate in the same direction; when negative, the fields counter-propagate. If the applied field is static ($\omega \rightarrow 0$), the cosine terms become unity, the time derivatives disappear, and one has the simpler coupled mode equations that are usually presented in the literature. (See, for example, Schmidt's article [7].)

Because the applied microwave field is monochromatic, one can use a harmonic analysis in which the field coefficients R and S are expressed as a Fourier expansion of forward-traveling optical waves. Thus,

$$R(z, t) = \sum_{m=-\infty}^{\infty} R_m(z) \exp[i m (\omega t - \beta_o z)] \quad (63)$$

with an identical expression for S . Because m is an integer, the coefficients R and S will be composed of harmonics of the microwave modulation frequency. Solving for all the R_m 's and S_m 's will thus give the complete optical spectrum, as well as the RF frequency spectrum of the demodulated (photodetected) lightwave.

After substituting equation (63) in equations (61) and (62), and equating terms with like time dependence, one finds the following infinite set of equations:

$$\frac{dR_m}{dz} = +i\frac{\delta}{2}(R_{m-1}\exp[i(\beta_o - \beta)z] + R_{m+1}\exp[-i(\beta_o - \beta)z]) - i\kappa S_m \quad (64)$$

$$\frac{dS_m}{dz} = -i\frac{\delta}{2}(S_{m-1}\exp[i(\beta_o - \beta)z] + S_{m+1}\exp[-i(\beta_o - \beta)z]) - i\kappa R_m \quad (65)$$

These equations are for a single microwave tone. When two tones are applied, as is normally done when measuring intermodulation distortion, equation (63) must be modified to read

$$R(z,t) = \sum_{m,n=-\infty}^{\infty} R_{mn}(z) \exp[im(\omega_1 t - \beta_{o1}z) + in(\omega_2 t - \beta_{o2}z)] \quad (66)$$

If the two tones are of equal amplitude, and sufficiently close in frequency that $\beta_{o1} \cong \beta_{o2} = \beta_o$, one finds that

$$\frac{dR_{mn}}{dz} = +i\frac{\delta}{2}((R_{m-1,n} + R_{m,n-1})\exp[i(\beta_o - \beta)z] + (R_{m+1,n} + R_{m,n+1})\exp[-i(\beta_o - \beta)z]) - i\kappa S_{mn} \quad (67)$$

$$\frac{dS_{mn}}{dz} = -i\frac{\delta}{2}((S_{m-1,n} + S_{m,n-1})\exp[i(\beta_o - \beta)z] + (S_{m+1,n} + S_{m,n+1})\exp[-i(\beta_o - \beta)z]) - i\kappa R_{mn} \quad (68)$$

These equations can be solved numerically by truncating the expansion at some upper value of m and n , and by using the “NDSolve” algorithm in *Mathematica*, which can solve a set of 18 coupled differential equations along the entire length of the coupler in less than a second. More importantly, the code itself is just a paragraph long, and thus simple to write. Accuracy is checked by increasing the number of harmonics used until changes in the solved value for the highest harmonic or intermod of interest become insignificant. For the third order intermod, $m = n = 3$ is usually enough.

It is instructive to examine the various intermods as they propagate along the directional coupler. Figure 37 shows the absolute value of the amplitude of the third order intermod ($m = 2$, $n = -1$) as a function of distance along a one-centimeter directional coupler having $s = 3.0$ (the nonlinearized value). Note that the intermod grows, decreases to zero, and grows again. If one

were to cut this modulator so that if its length is 0.953 cm, the intermod would be zero. This is exactly what one does in setting $s = 2.861$.

Figure 37 illustrates one of the primary problems in manufacturing the Dual-Y directional coupler modulator. If the length is not $2.861 l_o$, suppression of the intermod is incomplete. In fact, if one misses the optimum length by 5%, in either direction, there is no suppression at all. On the other hand, if one is within 0.5 % of the correct length, the intermod amplitude is one-tenth the non-suppressed value, corresponding to a 20 dB suppression of intermod power. Although this sounds impressive, a 20 dB suppression will only give a 6.7 dB improvement in the SFDR (20 divided by 3) because of the cubic dependence on drive power. Nevertheless, this might be adequate for applications requiring only a modest improvement in SFDR.

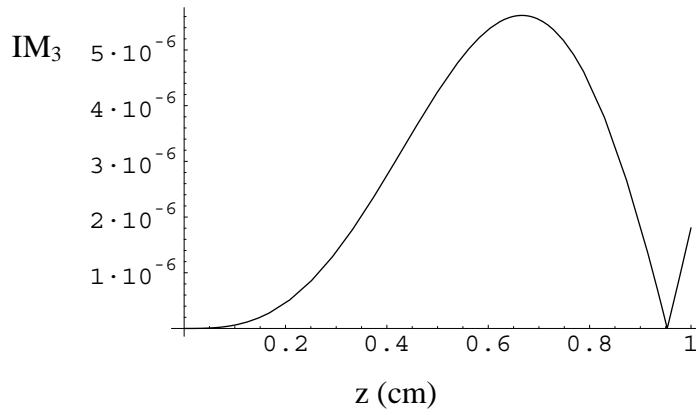


Figure 37. The amplitude of IM3 as a function of distance along the coupler for $s = 3.0$.

It should be emphasized again that the problem is not achievement of the correct *physical length* (which can be done with an accuracy of better than one micron). The problem is achieving the correct *coupling length*, which depends on κ . For material systems such as LiNbO_3 , which use titanium in-diffusion to form the waveguides, holding a 0.5 % tolerance on κ is very difficult.

4.3.3.2 Microwave and Lightwave Velocity Mismatch

When the microwave and optical waves travel at different velocities, an effect similar to that shown in Figure 37 occurs: the coupler is perceived to have the wrong length, so that the intermod is no longer suppressed. A good benchmark for back-of-the-envelope error analysis is the degree of velocity mismatch that raises the intermod to its non-suppressed value. Our analyses have shown that this occurs for IM_3 when

$$\frac{2\pi f \Delta n L}{c} \cong 1.1$$

where Δn is the difference between the microwave and optical index of refraction, f is the modulation frequency, and c is the velocity of light. Thus, a centimeter-long modulator operating at 10 GHz could suffer a 23 % index mismatch before losing all suppression.

However, a 4 cm long device (more typical of today's LiNbO₃ devices) could tolerate only a 5.9% mismatch. Reducing this to 0.59 % would give a SFDR improvement of 6.7 dB, reducing it to 0.059% would give 13.3 dB, and so on.

This improvement with improved manufacturing accuracy does not go on forever. At some point the intermods fall beneath the noise floor; when this happens, further suppression produces no further increase in SFDR. Because of this, the amount of SFDR that one wins with any linearization scheme is reduced as the noise bandwidth is increased. However, because there is less to win, the higher the noise floor, the less sensitive the device is to manufacturing errors. This is illustrated in the figures to the right, which show the SFDR of the Dual-Y directional coupler as a function of $\omega\Delta n L/c$ for three different noise bandwidths.

One sees that the SFDR values for the three different bandwidths are lower than those for the Dual-Parallel and Cascade Mach Zehnders. This difference is due to the fact that the Dual-Y Directional Coupler modulator is fundamentally less linear than a Mach Zehnder modulator. An examination of the transfer functions for the non-linearized directional coupler and a simple Mach Zehnder shows that the ratio of the cubic to linear term for the directional coupler is three times larger than that for the Mach Zehnder. A little analysis shows that the SFDR for the Mach Zehnder should thus be $20/3 \log_{10} 3$, or 3.18 dB better than the Dual-Y directional coupler. This is indeed what we find. This poorer linearity also affects the linearized case: for a 1 Hz noise bandwidth, the SFDR of the linearized directional coupler is 3 to 9 dB less than that for the linearized Mach Zehnder modulators.

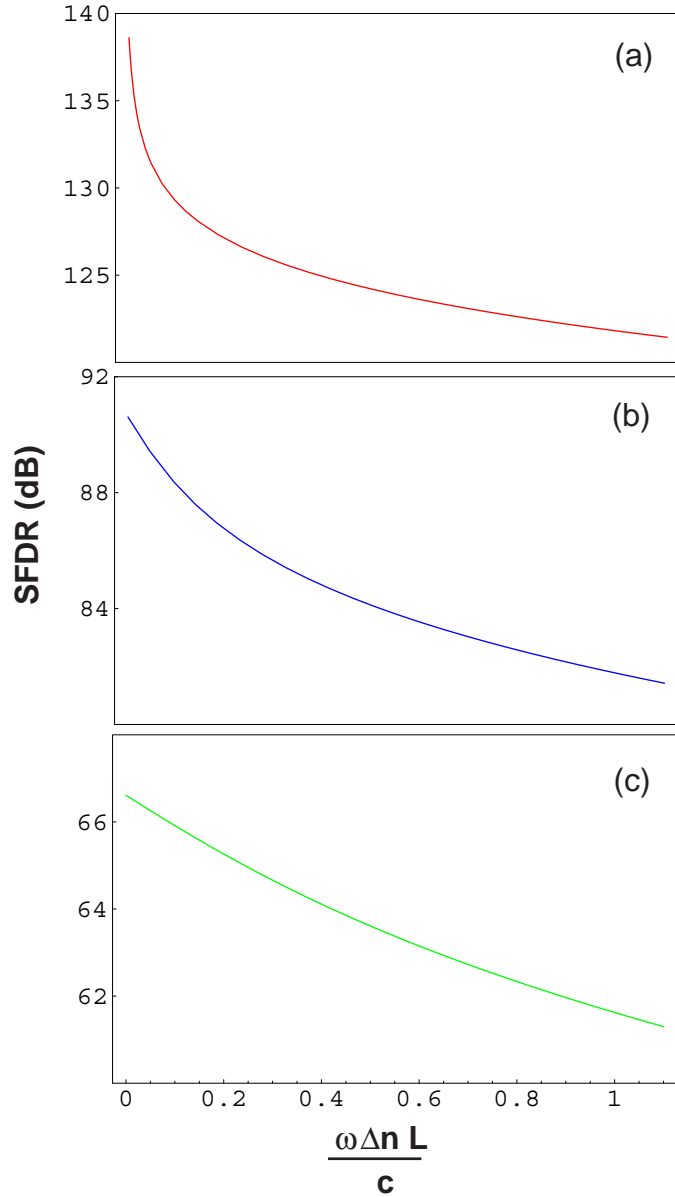


Figure 38. SFDR for three different noise bandwidths (a) 1 Hz, (b) 1 MHz, and (c) 1 GHz., as a function of velocity mismatch.

4.3.4 Electrode Attenuation

The Dual-Y Directional Coupler Modulator is almost 3 coupling lengths long. The first coupling length generates a modulated lightwave. The second coupling length undoes this modulation, and the third regenerates it. The lightwave that comes out of a 3-coupling-length device thus has the same depth of modulation as that from a single-coupling-length. The only function of the last two coupling lengths in our device is to provide linearization.

The synthesized devices discussed in section 4.2 operate similarly. The corrector sections that are added at the beginning and the end of the device do not contribute to the depth of modulation. In fact, they often reduce it. Higher degrees of linearization require additional corrector sections that add only to the length, and nothing to the depth of modulation. Thus, linearized directional coupler modulators are always longer, and usually much longer, than nonlinearized directional coupler modulators having the same depth of modulation.

Other linearization schemes that use two or more modulators in series or in parallel [9] operate by using a secondary modulator to generate a distortional component of the same magnitude but opposite sign of that generated in the primary modulator. When the signals are combined, the two distortional components cancel, thus providing a degree of linearization. The two directional coupler linearization approaches discussed in this final report do this too. However, it is not always obvious which section is the primary modulator, and which is the secondary. For the synthesized approach, it is fairly clear that the short corrector sections at the beginning and end of the modulator are the source of the distortional component, and are thus the secondary modulators. For the DYDCM, however, it is not so clear, since each section provides an equal depth of modulation (although of alternating sign). Nevertheless, for the discussion that follows, we'll think of the first coupling length as being the primary modulator, and the two remaining coupling lengths as being the corrector modulator.

The analysis performed up to this point has assumed that the applied voltage was uniform across the entire length of the device. If that is not the case, the voltage at the corrector section will be larger or smaller than that at the primary section. The distortional signals generated in the two different sections will then no longer be equal, perfect cancellation will not occur, and the SDFR will be reduced.

The resistive loss of the traveling-wave electrodes can have just this effect. The ohmic loss causes the microwave signal to decay exponentially as it travels forward along the device, so the applied field at the end of the device (the corrective section) will be smaller than that at the front (the primary section). Because of skin-depth effects, this loss increases with increasing modulation frequency.

To calculate this effect, we have extended the analysis of the previous section by including the exponential decay of the field. Equation (61) and (62) were modified to read

$$\frac{\partial R}{\partial z} + \frac{n_o}{c} \frac{\partial R}{\partial t} = +i \delta \exp[-\alpha z] \cos[\omega t - \beta z] R - i \kappa S \quad (69)$$

$$\frac{\partial S}{\partial z} + \frac{n_o}{c} \frac{\partial S}{\partial t} = -i \delta \exp[-\alpha z] \cos[\omega t - \beta z] S - i \kappa R \quad (70)$$

and were solved using the techniques described earlier. The frequency-dependent attenuation constant, $\alpha(f)$ was determined from measurements made on Y-Fed IMDD directional coupler modulators.

Figure 39 shows the reduction in SFDR that occurs for a device having the exact length needed ($s = 2.86059$) for complete removal of the 3rd order distortion term. The value at DC is 138 dB, as expected. At higher frequencies, however, the fall-off is precipitous. The bandwidth over which the SFDR is greater than 130 dB is slightly less than 1 GHz.

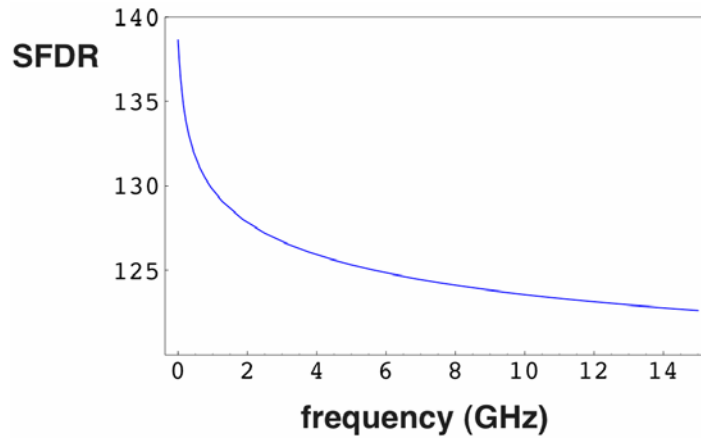


Figure 39. SFDR, for a 1 Hz bandwidth, as a function of modulation frequency for a DYDCM with $s = 2.86059$.

An interesting effect occurs, however, if s is less than the magic value of 2.86059: *the point of maximum SFDR shifts from DC to a higher frequency*. This is shown in Figure 40 for a DYDCM with an s of 2.810. One sees that the peak has shifted up to about 6.5 GHz. Using our simple model, one can argue that this occurs because the corrector's distortional term is too large at DC, but becomes smaller at higher frequencies because of electrode attenuation. Thus, at a particular frequency, the primary and secondary distortional terms become equal, and complete cancellation occurs.

In actuality, this is not quite correct. The magnitude of the combined primary and corrector distortional signals is still relatively larger, but the component that is selected by the homodyne detection process does indeed go to zero. (The homodyne detection process selects components that are in-phase, and rejects those that are 90° out-of-phase.) Thus, achieving the high SFDR values shown in Figure 40 would require an increased accuracy in the phase-locking circuit of the receiver. For SFDR values greater than 130 dB, the phase-locking error and phase fluctuations would have to be less than 6°.

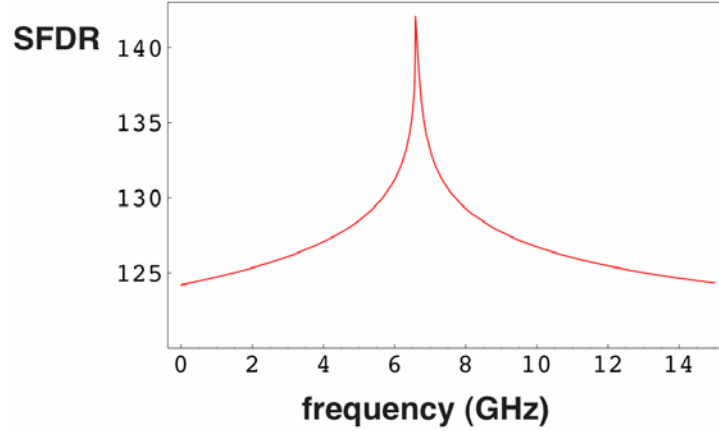


Figure 40. SFDR, for a 1 Hz bandwidth, for a DYDCM with $s = 2.810$.

From an experimentalist's point of view, the options offered by Figure 40 are intriguing. The plot implies that one could take a modulator that had been manufactured with the wrong value of s , and demonstrate a very high SFDR at a particular frequency that one could find by simply doing a frequency sweep. Thus, it would not be necessary to do any electrode etching or physical cutting of the device to get the right operating condition; one would simply tune to the relevant frequency.

There are also system implications for narrowband links. If the system need was to operate only over a small bandwidth, one could fabricate the devices with the value of s that would shift the peak to the center of that particular band. A SFDR of > 130 dB would be available over a 1 GHz bandwidth.

A final comment about Figure 40. One sees that the maximum SFDR shown is somewhat greater than the 138 dB value expected. This occurs because for all linearization schemes there is an operational point where one can actually achieve a SFDR value that is several dB larger than that achieved by a simple removal of the 3rd order distortional term. However, operation about this point leads to “cliffs” in the SFDR map [9] that are more confusing than useful, so we have avoided this issue by designing for a removal of the third order term only. In Figure 40, the use of a sub-optimum value of s and the sweep in frequency allows one to pass through this operating point, and pick up the few extra dB of SFDR that are available.

5.0 DEVICE MEASUREMENT

5.1 The PACT Program Modulators

The analyses of the previous sections have shown that the Dual-Y Directional Coupler Modulator should achieve a high degree of linearization if correctly fabricated. However, there were reasons for concern. Over the years, a large number of the linearization schemes that were reduced to practice have been found wanting. The devices often were either unstable, or required a degree of precision in control of some parameter that was not realistic, or simply didn't work because of manufacturing problems. We were thus apprehensive that we could spend considerable time and effort making these devices, and then find that they didn't work for reasons that had not been anticipated.

We were thus delighted to be able to take advantage of the fact that there was another program at HRL developing a similar device, although for an entirely different application. The PACT program [10] was designing and fabricating Y-fed IMDD directional coupler modulators for use in an optical A-to-D converter. These devices were almost identical to the Dual-Y Directional Coupler Modulators that we were developing, but differed in that they did not have the second Y-combiner nor the requisite 180° phase shifter. The PACT device is shown in Figure 41, the transfer curve in Figure 42.

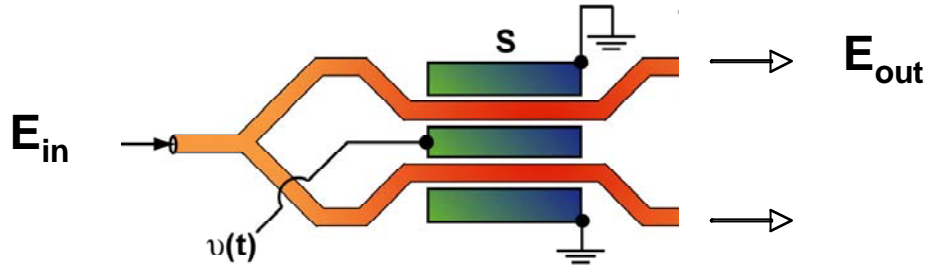


Figure 41. A Y-Fed Directional Coupler Intensity Modulator for IMDD applications.

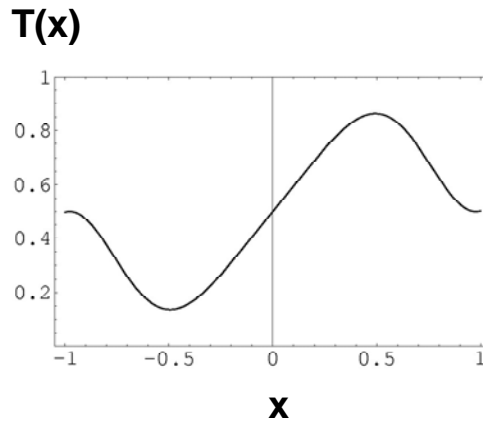


Figure 42. Intensity transfer curve for the modulator of Figure 41, with $s = 2.861$.

At our request the PACT Program Office graciously transferred the devices that had been fabricated on their program to this program, and allowed us to modify them as needed for our measurements.

The Y-fed directional coupler intensity modulator (YFDCM) can be linearized by making $s = 2.86059$, the exact same value needed for our coherent-AM device. The plan was to adjust the length of one of these devices to achieve the desired s , and to then measure the degree of linearization by monitoring either the 3rd harmonic or IM3 using the modulator as an intensity modulator (not an amplitude modulator), and detecting by direct detection. The use of Intensity Modulation and Direct Detection (IMDD) would greatly simplify the measurements, thereby allowing the expeditious screening and measurement of many devices.

If good linearization could be demonstrated, the plan was to then demonstrate a coherent device either by adding a fiber-based Y-Junction/Phase-shifter to the PACT device and measuring the device in a coherent AM link, or by fabricating a complete DYDCM in lithium niobate. The path chosen would depend on the available time and resources.

The fiber-based Y-Junction/Phase-shifter mentioned above is shown in Figure 43. It consists of a 3-dB coupler and two lengths of fiber, one of which would either be heated or stretched in order to give a controlled 90° phase shift. The assembly would then be spliced into the AM link of Figure 6, replacing the MZ modulator.

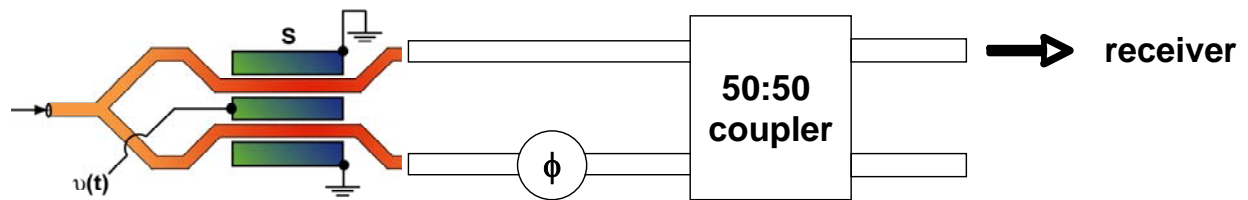


Figure 43. A method for using two fibers, a phase-shifter and a 3-dB coupler to mimic the Y-junction/Phase-shifter needed for a DYDCM.

5.2 Length Adjustment

As pointed out earlier, one cannot change s on a Y-fed Directional Coupler *Intensity* Modulator by changing the electrode length. One must physically change the length by sawing through the modulator and discarding the unwanted portion. Two things happened when we did this. The first was that we cut off the probe pads at the end of the modulator, so that we could no longer make RF contact at the termination end of the device. We were aware of this, and had purchased a special probe from GGB Industries (Picoprobe) that had a 14 micron tip-to-tip separation which would allow us to touch down on the two long 50-Ohm Co-Planar-Strip (CPS) electrodes that covered the waveguides. Thus, we could launch the drive signal from the uncut end, which still had a launch pad, and use the special probe and a 50-Ω termination attached to the probe to terminate the CPS line.

The second thing that happened was that the sawing operation, which used a 30,000 rpm high-speed diamond blade with a water jet, moved some of the soft gold of the electrode from one waveguide electrode to the other, electrically shorting the two electrodes together. It was first believed that the backward wave generated on the CPS by a reflection at this short would not affect the performance of the modulator. However, that proved not to be the case. Measurements and analyses showed that the linearization properties of the YFDCM were indeed affected by the reflected wave. Thus, a way had to be found to open up these shorts.

Mechanical techniques were out of the question because of the small feature size (~ 10 microns): a micro-scribe would merely increase the damage. It was decided to remove the smeared gold at the edge of the cut by chemically etching it away. The first attempts to do this by simply placing the wafer vertically in a very thin layer of etchant proved unsuccessful. Capillary action wicked the etchant solution up the electrodes, removing much more gold than was desired. It was then decided to cover the wafer with resist, and lithographically expose the launch pads while leaving the waveguide electrodes covered. The wafer would then be sawed while the resist was on, and the troublesome gold removed chemically.

This technique, after several iterations, seemed to work quite well. Figure 44 shows two different electrode pairs after sawing and after chemical removal of the gold. The gray strips that remain after the gold removal for the device on the left are due to the 100-Angstrom-thick layer of titanium used for adhesion of the gold to the 1-micron-thick SiO_2 buffer layer covering the LiNbO_3 . In a few cases the ends of the gold electrodes were bent to the left by the blade, as shown in the figure on the right. However, even these devices were electrically open after etching.

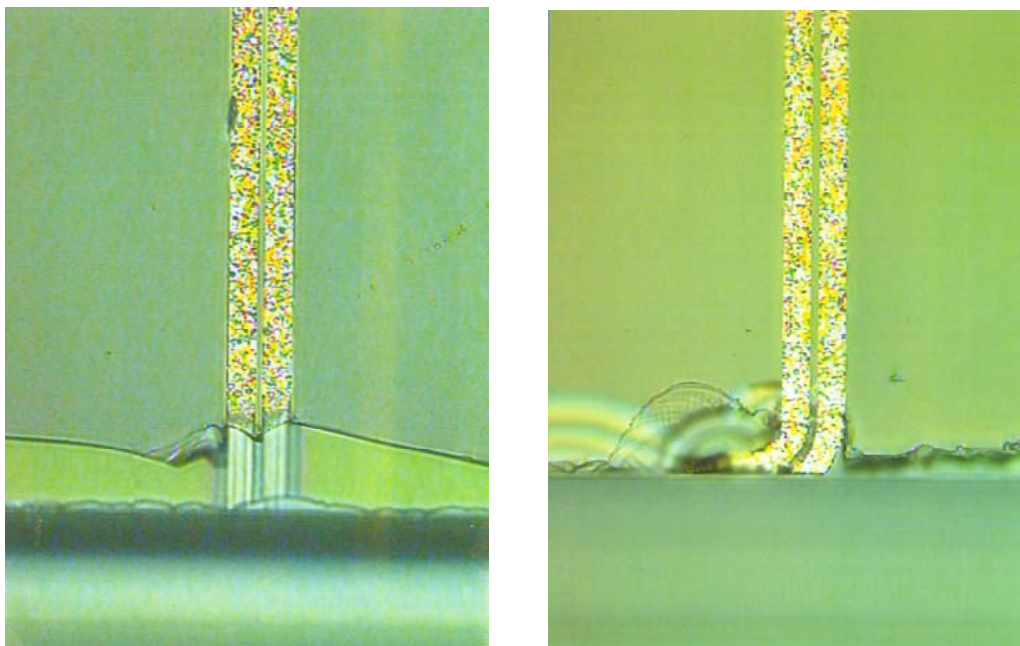


Figure 44. The ends of the YFDCM electrodes after etching. The devices are still covered with a thin layer of resist. The bending of the electrodes of the device on the right did not cause shorting.

Figure 45 shows the probe pad at the beginning of the modulator. The pad, which is designed for either a 100 or 150-micron-pitch probe, has had the resist removed lithographically so that DC and RF contact to the pad can be made.

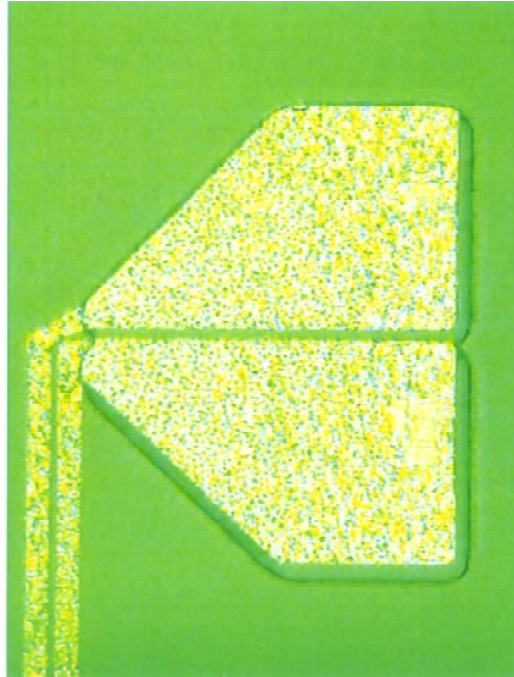


Figure 45. Launch pad at the top of the modulator. The resist above the pads has been removed lithographically so that RF contact can be made.

Our approach for achieving the correct coupling length, s , was to now do the following:

1. Make an initial cut at a point where s is *known* to be > 2.861
2. Remove the short by locally etching away the gold
3. Measure the 3rd harmonic as a function of drive power. Determine position of next cut
4. If logarithmic slope is exactly 5, stop.

Before doing this, however, we went through a screening procedure to select those wafers that had the best chance of success.

5.3 Screening

There were a large number of wafers that had been generated on the PACT program, but only a few were appropriate for our program. We thus had to institute a screening procedure to select wafers that had devices with the appropriate length, that were not internally shorted and that had transfer curves that followed the theory (indicative of good devices). The devices had to be calibrated so that not only the approximate value of s but also the calibration factor for x (the relation between applied voltage and the x of equation (49)) was known for each. Calibration of x was a tedious process that we will describe shortly, and because of this was only done once per wafer and only on those wafers that looked really promising. To cull out those wafers that were unsatisfactory, we used the large amount of transfer curve data that had been generated on the PACT program, provided to us courtesy of the PACT Program Office. The data for one of the selected devices are shown in the Figure 46.

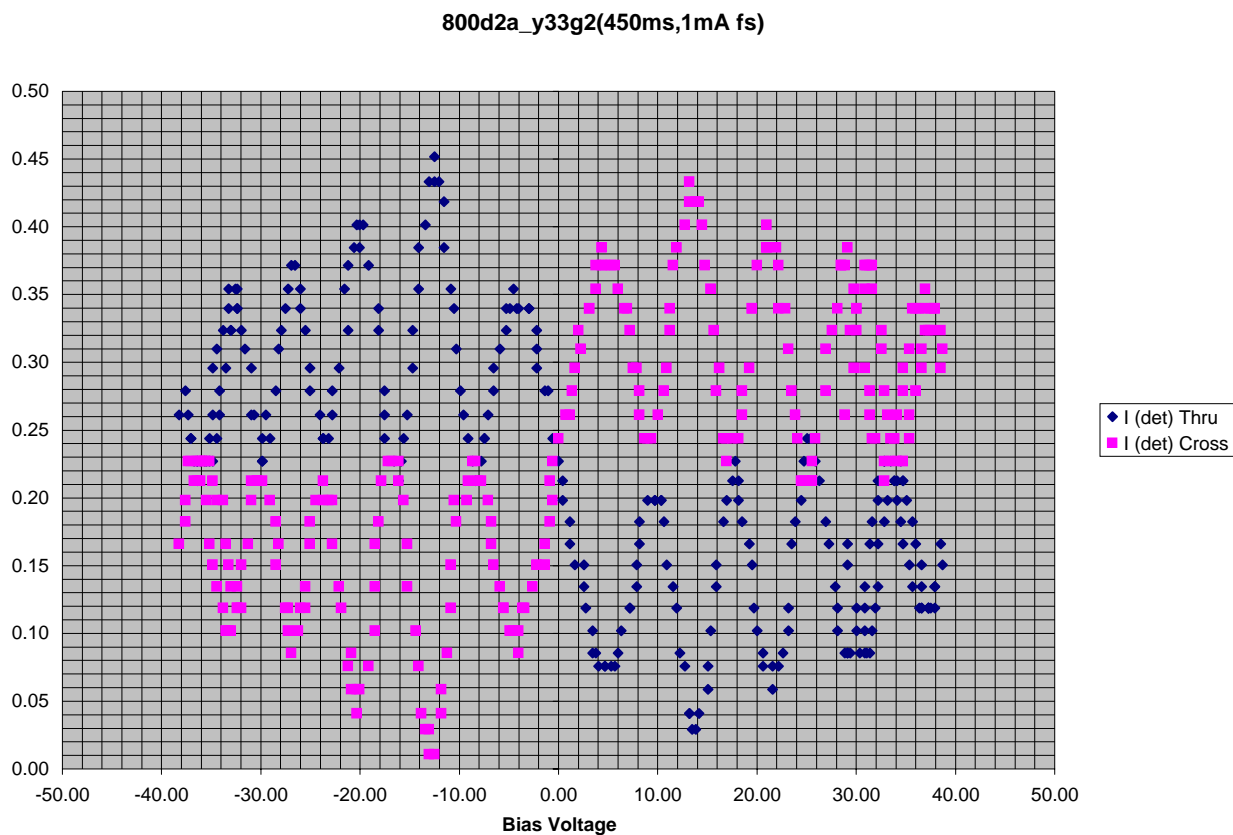


Figure 46. Measured transfer curve for a YFDCM. The two colors are for the two different output ports.

5.4 Harmonic Measurements

We have developed a technique that allows one to completely characterize all of the operational parameters of a Y-fed directional coupler modulator operating as an IMDD device. The accuracy for x and s is approximately 2 and 3 significant figures respectively, which though not quite good enough to determine s to the four significant figures needed, is more than enough to calibrate x , which then allows a more accurate determination of s through another measurement that we'll describe later.

The technique is as follows. The modulator is mounted in our test setup, 1.55 μm light is fiber-coupled into the launch end of the modulator, and the output light is fiber-coupled with a lensed-fiber into a calibrated photodetector that is part of a lightwave measurement system (Figure 47). An RF signal at 145 MHz is first passed through a bias-T, then through a narrow filter to remove any spurious harmonics, and then finally launched onto one end of the modulator using a 100-micron-pitch Signal-Ground Picoprobe. A DC bias voltage is applied to the modulator electrodes through the bias-T, and the fundamental, 2nd and 3rd harmonics generated in the photodetector measured on the spectrum analyzer.

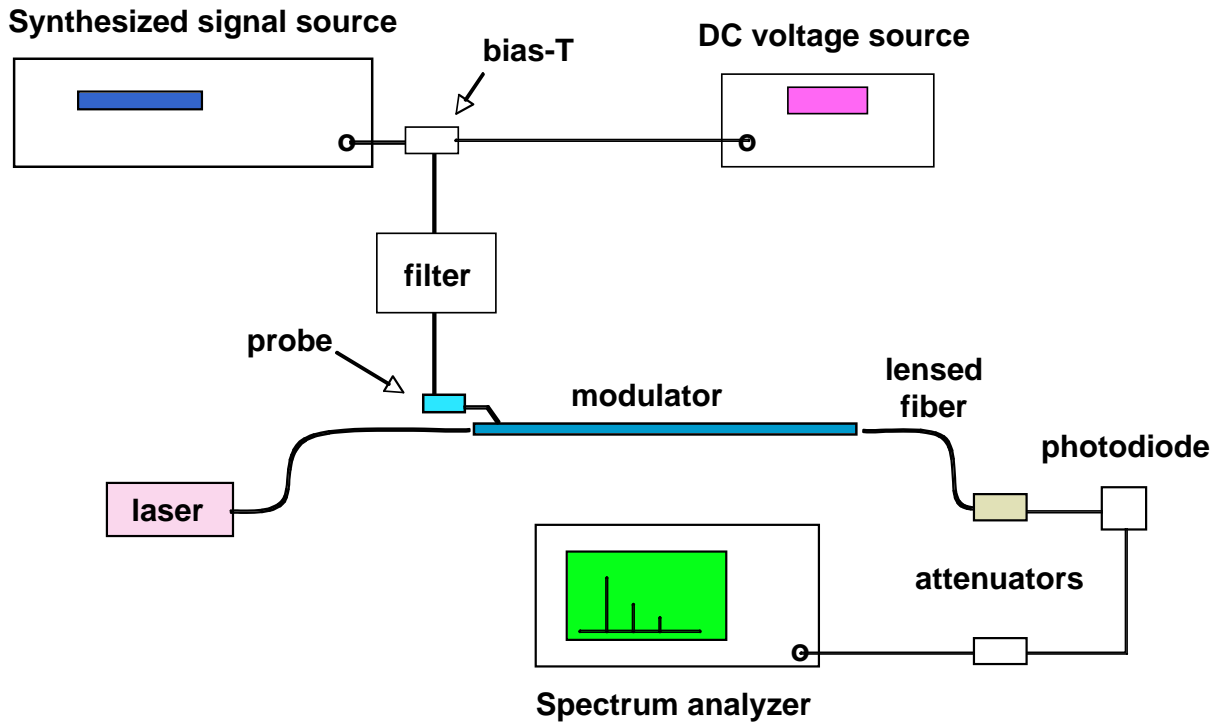


Figure 47. Harmonic measurement system.

The lightwave measurement system has switchable optical and electrical attenuators in front of the photodetector and microwave spectrum analyzer, respectively. This allows one to rapidly switch attenuators in and out to determine whether photodetector or spectrum-analyzer nonlinearity is interfering with the measurement.

The measurement is performed by measuring the 1st, 2nd and 3rd harmonics that are generated for a fixed-frequency and fixed power level signal at a particular DC bias voltage. The bias voltage is then changed, the system allowed to equilibrate, and the measurement repeated over a 15 to 30 volt range centered at zero. The results of one of these measurements is shown in Figure 48. The blue points are the measured values the black curves are the theoretical fit. With the exception of the transfer curve, the vertical scales are all power in dBm.

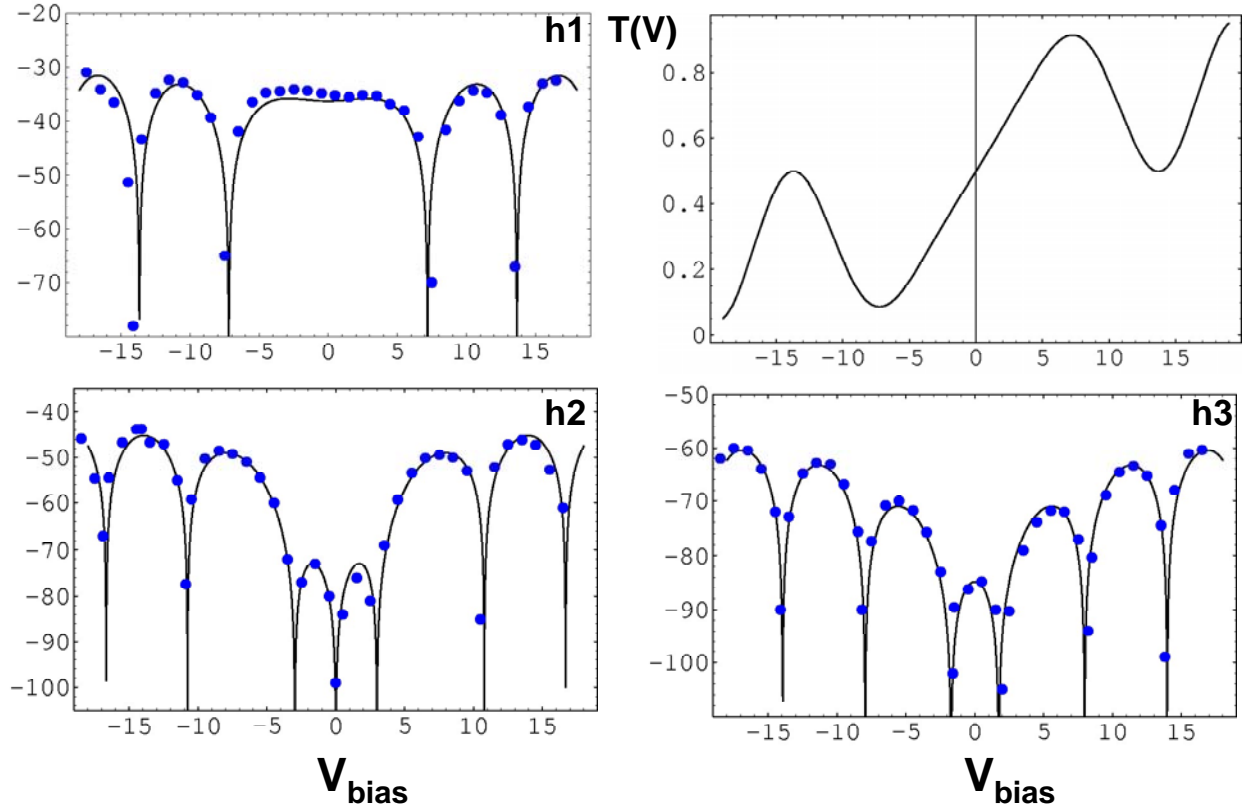


Figure 48. The 1st, 2nd and 3rd harmonic as a function of DC bias voltage on the modulator.

The high accuracy of the measurement is due to the fact that there are only three adjustable parameters, s and x and the DC switching voltage that one can adjust to fit all the data. A small error in any one of these leads immediately to a poor fit. The fit values for this particular curve are: $s = 2.72$, $x = 0.13$ and $V_s = 12.7$ Volts. This particular modulator is number Y30G4A from wafer 750N4. The total optical power coming out of the modulator is 0 dBm (1 milliwatt), the RF drive power at the probe is 10 dBm. Because the modulator was driven at a low frequency, the CPS electrodes did not have to be terminated.

Using the transfer curve data available to us, we did a search of the 25 existing wafers, and found 4 that were appropriate. On these 4 wafers, we found 4 modulators that were close to $s = 2.86$. We decided to measure these 4 devices carefully, before doing any cutting, on the off-chance that one already had the magic value of $s = 2.861$.

The measurement that most accurately determines s , and that allows us to accurately predict the value of SFDR that that particular device would deliver, is just a measurement of the 3rd harmonic power versus drive power. Figure 49 shows a typical measurement of one of the four “hot” devices. The blue curve is a theoretical fit to the black data points that gives the value of s shown (2.873). Hence, this device is only 0.4% longer than the magic value of 2.861. The slope for this device is 4.4, not 5, but 4.4 already denotes a significant amount of linearization. However, it is not enough to get us to our goal of a SFDR of 130 dB in a 1 Hertz bandwidth.

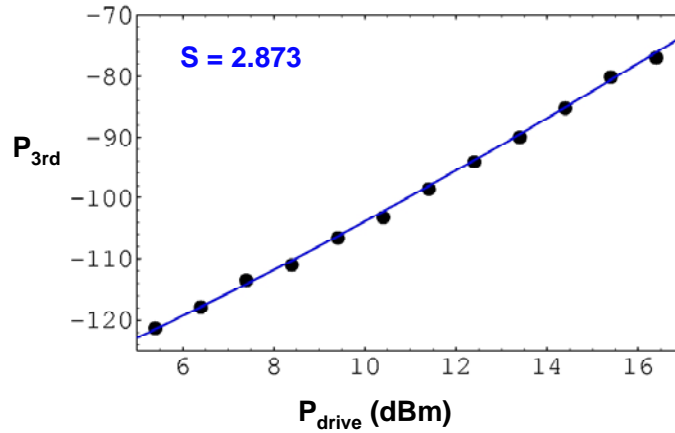


Figure 49. 3rd harmonic power vs drive level for device Y31G4B, wafer 750N4.

We then made several measurements on devices that we were sure had a value of s less than 2.861. When s is less than 2.861, the 3rd harmonic curve should show a sharp null (a trough) at some power level. However, our measurement did *not* show the null predicted by theory, as is evident in Figure 50. This was very disturbing. In addition, the measurement was occasionally very noisy when operating where the trough should be, with the data points jumping up and down by 10 dB.

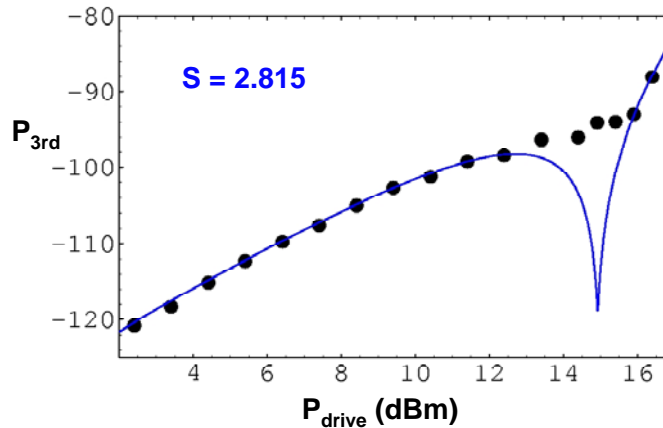


Figure 50. 3rd harmonic power vs drive level for device Y30G4B, wafer 750N4.

Our assumption at the time was that we were having problems with bias drift. This is a common problem in LiNbO_3 , and is related to the fact that the SiO_2 coating that one puts on top of the wafer to isolate the optical waveguides from the metal electrodes to prevent optical loss is less conductive than the LiNbO_3 underneath. When this is the case, the application of a voltage to the electrodes will generate an electric field in the LiNbO_3 that will, due to the larger conductivity of the LiNbO_3 , decay with time. After a long enough time, the field in the LiNbO_3 becomes zero with all of the field ending up in the SiO_2 buffer layer. This is shown figuratively in Figure 51.

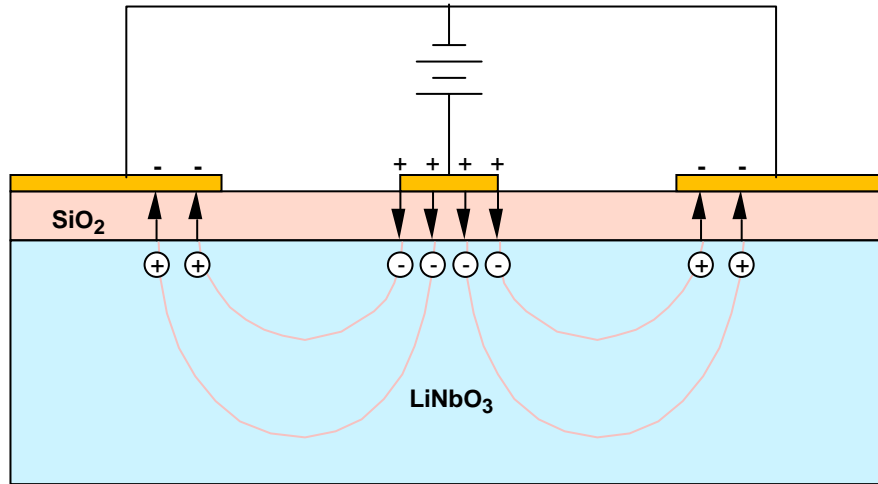


Figure 51. The electric fields in a LiNbO_3 modulator when bias drift is a problem. With time, the dashed field lines disappear, and all of the field ends up in the SiO_2 buffer layer.

If the field decays with time, then the actual bias point of the modulator changes. Our thinking was that if the LiNbO_3 conductivity varied along the device, then different regions of the modulator would see a different bias voltage, and this could ruin the linearization.

The known fix for this problem is to make the SiO_2 buffer layer more conductive than the LiNbO_3 . This can be done with a special coating process that had been developed on another program. And we were in luck (or so we thought), because one of the good wafers had already had this coating applied.

We made some preliminary measurements on this wafer, and found a device that seemed to have a value of s even closer to 2.861. A complete harmonic scan with bias voltage was then performed (Figure 52). The device seemed to take a long time to come to equilibrium at each new measurement point, but it did have a much lower switching voltage, which is what one would expect for a high-conductivity SiO_2 . A measurement of the 3rd harmonic vs drive power gave the curve shown in Figure 53. A fit to the data gave an s of 2.857, which was indeed closer to 2.861 than the device of Figure 49. Portions of the curve showed a slope greater than 5, but this was consistent with a device having a null somewhere off the left side of the graph.

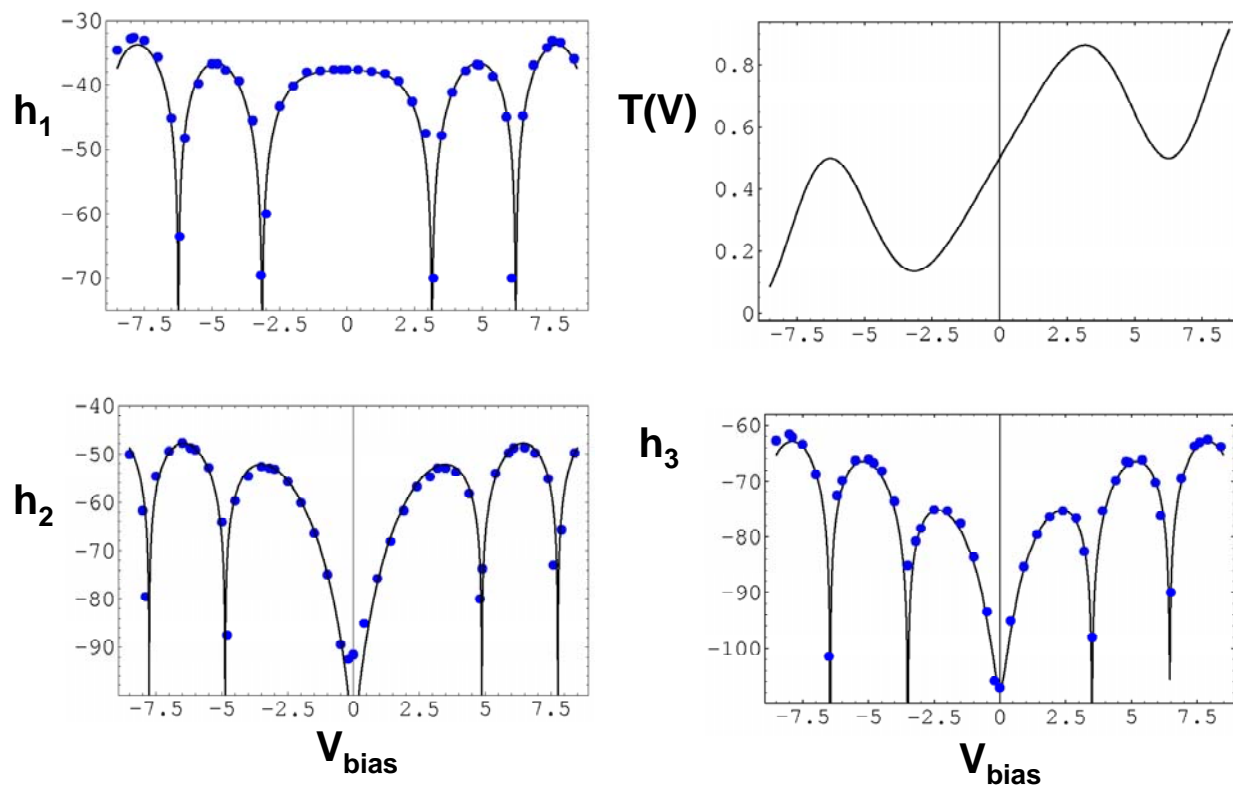


Figure 52. Harmonic data for device Y33G2A from wafer 800D. This wafer has the anti-drift coating.

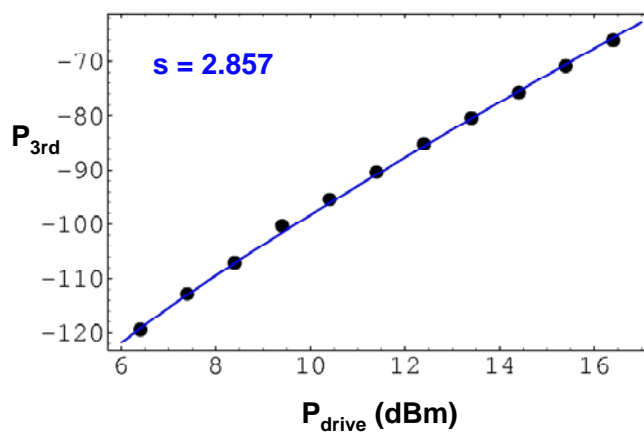


Figure 53. 3rd harmonic data for the device of Figure 52.

The big surprise came the next day when we measured the same device again, and got an entirely different curve, shown in Figure 54. It was now clear that we had a very serious problem of some sort.

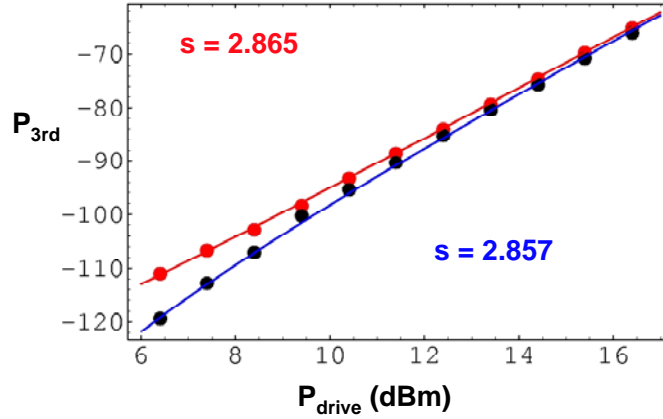


Figure 54. The device of Figure 53, measured a second time (the red curve).

The major change seemed to be at the lower end of the curve. To pursue this, we monitored the lowest data point over time, and found that it drifted slowly, sometimes going up, sometimes down. The drift times could be very slow (on the order of hours). We thus had three problems:

1. The data was not reproducible at lower drive levels
2. There was no real “trough” for devices with $s < 2.861$
3. The data was very noisy when operating where the trough should be.

A list of all the possible mechanisms that could explain these observations was put together. One by one, over a period of several days, each of the items on the list was investigated and eliminated either experimentally or by analysis. Until we came to the last item on the list: ambient temperature. And then we hit the jackpot.

We had done some calculations of temperature effects, and found that all of the perturbative mechanisms that we could think of had no effect on modulator performance. Nevertheless, the effects we were observing seemed to have a thermal character. To investigate this experimentally, we mounted the wafer on a special aluminum chuck that we used to measure high-temperature drift effects in LiNbO_3 modulators, and measured the 3rd harmonic while raising the temperature of the chuck at a rate of 0.1 degree C per minute. The results, shown in cartoon fashion in

Figure 55, showed a large and cyclic behavior. And as soon as we saw this, we knew what the problem was.

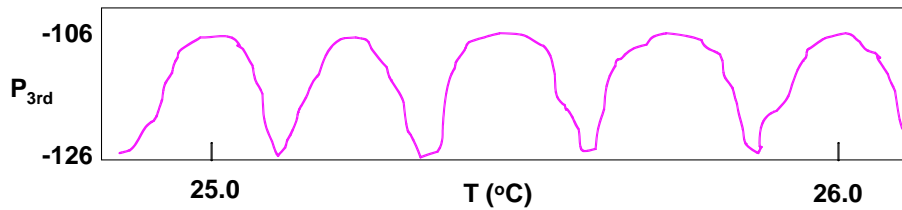


Figure 55. 3rd harmonic power as a function of device temperature for a fixed drive level.

5.5 The Fabry-Perot Modulator Effect

The source of the problem is a Fabry-Perot (FP) modulator effect. Because these devices are being measured on a full wafer, no bevel-cuts have yet been made to frustrate undesired reflections (this is normally done during packaging). Hence, the end-facets of the device being measured have two very nicely-polished surfaces at which small but non-negligible reflections can occur. This forms a Fabry-Perot etalon.

A Fabry-Perot etalon is a one-dimensional optical cavity or waveguide with partially-reflecting surfaces at each end. The etalon transmittance changes with wavelength due to multiple internal reflections that can add constructively or destructively depending on the etalon length. A simple waveguide etalon is shown in Figure 56

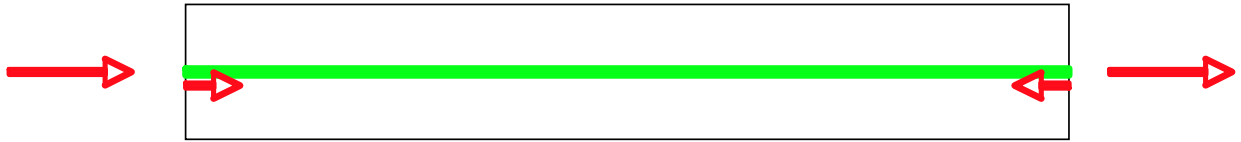


Figure 56. A Fabry-Perot etalon. The small red arrows inside represent the light reflected at each surface.

The transmittance of the etalon will also change if one varies the index of refraction of the waveguide material. The optical length of the cavity (kL) will then change, moving the operating point to a different location on the etalon transmittance curve shown in Figure 57. If the operating point is on one of the steep slopes of this curve, and the index is dithered at high frequency, one gets significant high-frequency modulation. If, on the other hand, the operating point is at the top or bottom of this curve, the modulation will be close to zero. If the operating point now moves to the opposite side of one of these bumps, then the modulation due to index dithering will again be significant, but of the opposite phase.

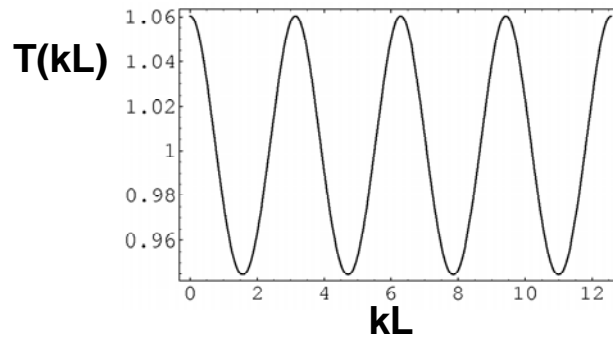


Figure 57. The transmittance of a LiNbO_3 /glass etalon as a function of optical length.

The dithering is provided by the microwave field and the electrooptic effect of the material. The change in the operating point, however, is provided by the thermally-induced index change and mechanical expansion of LiNbO_3 . Thus, as the temperature of the material slowly increases, one

slowly moves from left to right on the transmittance curve of Figure 57, passing through points of positive modulation, zero modulation, and negative modulation.

The modulation itself is not necessarily a problem. The nonlinearity of the etalon transmittance curve, however, is. This nonlinearity produces a 3rd harmonic that can either add to or subtract from the 3rd harmonic naturally produced by the YFDCM, depending on temperature. We have done a calculation of this effect, and have found that it has the right magnitude to explain the observed data. In addition, a calculation of the length and index change in LiNbO₃ showed that, for the length of this particular wafer, one would expect 5.4 cycles for a 1° C temperature change of the modulator. This is exactly what we see.

The question now is whether this can explain all of the nasty effects we have seen. And the answer is yes. Figure 58 shows two sets of data from the same device. If we assume that a FP modulator effect is generating a 3rd harmonic that is interfering with our measurement, then one can determine the magnitude of that harmonic from the maximum and minimum values of the 3rd harmonic measured for a particular drive level during the temperature sweep (in this case the lowest data point). By knowing the maximum value at one drive level, and by knowing that it follows a cubic power law (slope 3), one can deduce the 3rd harmonic power for all other drive powers. We have done this for the modulator of the left figure, showing the FP power that results as a red line. That same curve was then plotted in the figure to the right, which was a device from the same wafer (and should therefore have the same FP modulator effect). Note that the red line deduced from measurements on the left device explains the observed data for the device on the right. There is no trough on the right-hand figure because the FP modulator effect is “filling” the trough. Furthermore, the data in the trough will be very noisy whenever temperature puts the FP modulator operating point at either the top or bottom of the transmissivity curve: when this happens, the trough will empty because the red line has dropped way down; very small temperature fluctuations, however, will cause it to refill, then empty again, and so on. This explains the very large (10 dB) fluctuations that were occasionally observed.

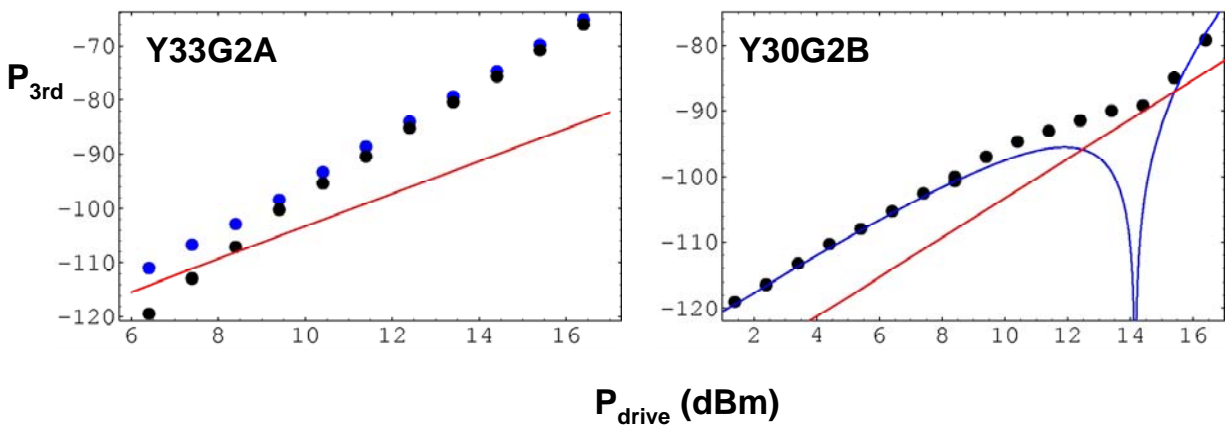


Figure 58. Measured 3rd harmonic data (points) and the FP modulator 3rd harmonic power (red line) for two devices from the same wafer. The blue and black points in the left figure are for measurements made at two different times. The blue curve in the right is the predicted response for a device with no FP effect.

Although these arguments are compelling, they do not prove that the FP modulator effect is responsible for the observed effects. To test this hypothesis, we measured another wafer before and after making a 6° saw cut at one end of the wafer. Experimental results for devices made on another program with this bevel have shown that the reflected wave is reduced by 33 to 35 dB, which is somewhat less than the > 40 dB predicted by our analysis, but a respectable reduction nonetheless. Because the fiber itself did not have a bevel, we reduced the reflection from the fiber-air interface by moving the fiber away from the output facet until the signal intensity had dropped by 10 dB, which should reduce the power reflected back into the modulator by 20 dB.

The before-and-after measurement data of Figure 59 show convincingly that the FP modulator effect was playing a strong role. The red data points (the “before” measurement) show a slight dip at the position of the trough. However when the FP effect was removed by the 6° cut, the trough deepened by more than 20 dB. Note the almost perfect fit of the data to the blue theoretical curve.

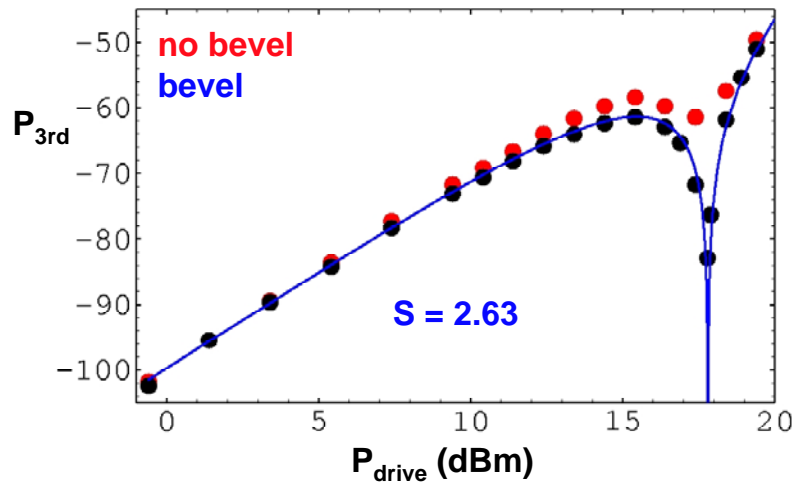


Figure 59. 3rd harmonic power for device Y34G3B, wafer 800D3. The black data points were taken after a 6° angle had been formed at the end of the wafer, the red data taken before.

The depth of the trough, though much improved, is not infinite. There is thus some other effect, 20 dB smaller than the FP modulator effect that is interfering. This is not surprising. As one travels down the slope 5 curve, one will invariably run into other effects having slope 3 that were too small to be seen in earlier measurements because they were always masked by the distortional tones of the modulator. This is shown in Figure 60. When the modulator is fixed somewhat (i.e., give a slope of 5), then at some lower level some “other” effect, (e.g., effect *a*), will manifest itself. And when effect *a* is fixed, effect *b* will surface, and so on. Whether or not these affect the SFDR depends on the position of the noise floor.

One can correct for the FP modulator effect rigorously by doing a temperature sweep and a correction at each data point. This, however, is very time consuming. What we did instead is take a data set in which it was clear that the FP effect was *subtracting* from the true level (Figure 53), and added the FP power to that data. This gave the data points and fit shown in Figure 61.

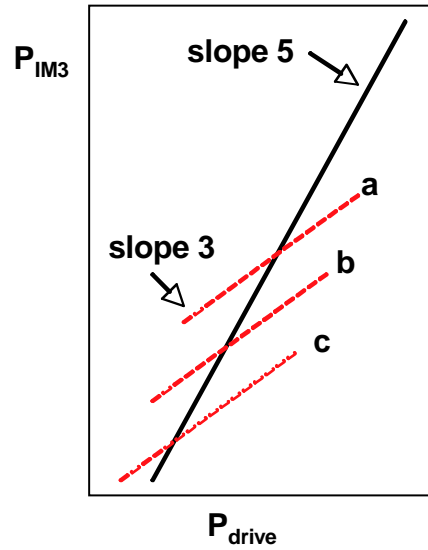


Figure 60. A plot showing how slope-3 effects will always manifest themselves when the primary distortional term is slope 5.

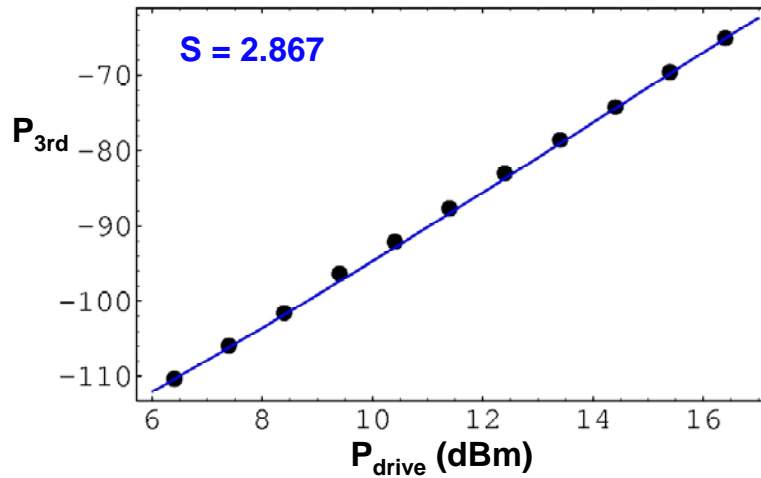


Figure 61. 3rd harmonic power for device Y33G2A, wafer 800D2, corrected to remove the Fabry-Perot modulator effect.

The slope of this curve is 4.6, which is still not the desired 5.0. However, a slope of 4.6 is not bad. The question is, is it good enough?

Table 5 shows the theoretical SFDR that one would achieve for a perfect device with $s = 2.8606$, and for our device having an s of 2.867. One sees that the SFDR for a 1-Hz bandwidth is 130 dB, which was our program goal. The table also shows how the advantage of linearization

falls off as the bandwidth increases, even for a perfect device. This fall-off is typical of all linearization schemes that eliminate the 3rd order distortional term.

Table 5. SFDR versus measurement bandwidth for two different values of s .

Bandwidth	$s = 2.8606$	$s = 2.867$	Mach-Zehnder
1 Hz	138.9	130.3	124.7
1 kHz	114.7	110.1	104.6
1 MHz	90.6	88.9	84.6
1 GHz	66.6	66.2	64.6

At the conclusion of these last measurements the program was almost over. The funding that remained at this point, roughly two-man-weeks, was used to chase down a serious error in the unlinearized coherent AM link SFDR measurement. This error was found and corrected, and a completely new set of measurements made to replace the data that were suspect. These new data were presented in this final report.

6.0 Summary and Conclusions

This program has shown that coherent AM optical links can be made that have a SFDR much higher than that available with IMDD links. Values of $120 \text{ dB Hz}^{2/3}$ were obtained for an unlinearized link using a 100 mW laser and COTS modulators and photodetectors. Slightly higher values ($123 \text{ dB Hz}^{2/3}$) were obtained when the laser power was boosted to 500 mW with a fiber optical amplifier. The link, which used polarization-maintaining components and fiber throughout, and fusion-splices instead of connectors, had none of the instabilities historically associated with homodyne systems. The system locked, and held lock, for the duration of all measurements (which often lasted all day). We saw no damage effects in the modulators used, even at the 500 mW power levels used at the program's end.

The plan to raise the $120\text{-}123 \text{ dB Hz}^{2/3}$ SFDR values demonstrated on the first part of this program to $130 \text{ dB Hz}^{4/5}$ by using a linearized directional coupler modulator was pursued, but was not completed. Y-fed directional coupler modulators having the correct configuration were loaned to us by the PACT program office, and an extensive set of measurements was conducted on these devices. After resolving a very serious measurement problem, it was found that the devices did indeed perform as theoretically predicted at low frequencies. The modulators were stable, did not drift once equilibrium was reached, and showed the linearization predicted by theory. A logarithmic slope of 4.6 was demonstrated for the best device; the only reason a perfect slope-5 was not demonstrated was because the 3.3-cm device tested was 66 microns too long, and not because it had any inherent flaws.

Our analysis showed, however, that Y-fed linearized directional coupler modulators (and perhaps all linearized directional coupler modulators) have a shortcoming that makes them unsuitable for wideband operation, at least for coherent-AM applications: RF attenuation in the CPS traveling-wave electrodes degrades the linearization at higher frequencies. The present devices can be designed to operate at a specific center frequency, but the linearization bandwidth at this frequency is less than 2 GHz.

Our measurements on the first part of the program also suggested that the photodetectors were beginning to limit the SFDR of the unlinearized link. Although this will have to be looked into more rigorously before any definitive statement can be made, the possibility exists that today's photodetectors could limit the performance of optical AM links to $\sim 127 \text{ dB Hz}^{2/3}$.

References

- [1] R. R. Hayes and D. L. Persechini, "Nonlinearity of p-i-n Photodetectors," *IEEE Photonics Technology Letters*, vol. 5, pp. 70-72, 1993.
- [2] R. R. Hayes, J. F. Lam, W. W. Ng, J. H. Schaffner, W. B. Bridges, and D. L. Persechini, "Fiber Optic Link," Hughes Research Laboratories, Malibu, CA 90265 91F135000, December 1992 1992.
- [3] C. Laliew, S. W. Lovsetg, X. Zhang, and A. Gopinath, "A Linearized Optical Directional-Coupler Modulator at 1.3 micron," *J. Lightwave Technology*, vol. 18, pp. 1244-1249, 2000.
- [4] S. W. Lovseth, C. Laliew, and A. Gopinath, "Synthesis of Amplitude Response of Optical Directional Coupler Modulators," presented at 1997 IEEE MTT-S Int. Microwave Symp. Dig., 1997.
- [5] R. F. Tavlykaev and R. V. Ramaswamy, "Highly Linear Y-Fed Directional Coupler Modulator with Low Intermodulation Distortion," *J. Lightwave Tech.*, vol. 17, pp. 282-291, 1999.
- [6] B. Desormiere and C. Maerfeld, "An Integrated Optic Frequency Translator for Microwave Lightwave Systems," *J. Lightwave Tech.*, vol. 8, pp. 506-513, 1990.
- [7] R. V. Schmidt, "Integrated Optics Switches and Modulators," in *Integrated Optics - Physics and Applications*, vol. 91, *NATO ASI Series*, S. Martellucci and A. N. Chester, Eds. N.Y.: Plenum, 1981, pp. 181-210.
- [8] A. Yariv, *Quantum Electronics*, 3rd ed. N.Y.: Wiley, 1989.
- [9] W. B. Bridges and J. H. Schaffner, "Distortion in Linearized Electrooptic Modulators," *IEE Trans. on Microwave Theory and Techniques*, vol. 43, pp. 2184-2197, 1995.
- [10] Photonic A to D Converter Technology program, AFRL Rome/DARPA-MTO, Contract No. F30602-99-C-0022.

Appendix A: Laser Excess Noise (REN)

To see just what effect laser amplitude and phase noise has on link performance, we must redo the analysis of section 2.2 with laser fluctuations included in the field equations. Using the well-known narrow band approach, in which the noise is expressed as in-phase and quadrature components, the signal and local oscillator fields are now given by

$$E_s(t) = E_s \theta(t) \left(\cos(\omega_c t) + \eta_s^I(t) \cos(\omega_c t) + \eta_s^Q(t) \sin(\omega_c t) \right) \quad (71)$$

$$E_{lo}(t) = E_{lo} \left(\cos(\omega_c t + \phi) + \eta_{lo}^I(t) \cos(\omega_c t + \phi) + \eta_{lo}^Q(t) \sin(\omega_c t + \phi) \right) \quad (72)$$

The noise terms, η^I and η^Q , are slowly varying functions of time representing the in-phase (I) and quadrature (Q) components of the noise for each laser. Note that, as written, the η 's are unitless quantities that specify what *fraction* of the respective field is noise. Thus, their powers are expressed in dBc.

The analysis performed in section 2.2 is now repeated. One sends the two signals through the combiner, with the fields going to the cross ports being phase-shifted by 90° , and those not, not. The summed fields are then squared by each detector, yielding time-dependent detector currents. One then computes the current into the amplifier by subtracting one photodetector current from the other. After maximizing the signal term by setting $\sin\phi$ to 1, and $\cos\phi$ to zero, one finds that

$$I_{signal} = I_1 - I_2 = E_s E_{lo} \theta(t) \left(1 + \eta_{lo}^I(t) + \eta_s^I(t) + \eta_{lo}^I(t)\eta_s^I(t) + \eta_{lo}^Q(t)\eta_s^Q(t) \right) \quad (73)$$

The first term is the desired signal, the second and third are the in-phase noise of the local and signal lasers, and the fourth and fifth are the noise \otimes noise terms. The noise \otimes noise has a spectrum that is roughly twice as wide as the in-phase noise of either oscillator. However, it is also much, much smaller. We shall ignore it for now, and focus on the two dominant terms.

If one knows the noise spectral density for both lasers, one can compute the noise spectrum of the receiver at all frequencies and for all operating conditions. However, one need not have full knowledge of the spectral shape to appreciate several salient features of equation (73).

The first is that $\theta(t)$ modulates the noise as well as the carrier. The effect of this is to translate the noise spectrum up to the modulation frequency, and convert it to a double-sided spectrum, centered at the modulation frequency, as shown in Figure 62 below.

The second feature is that the magnitude of the upconverted spectrum is proportional to the power in the modulation tone. Thus, if the laser noise exceeds the shot noise, the SNR becomes constant, regardless of signal strength (!). This is not good for those applications that look for weak signals in the presence of strong ones. The strong signal will raise the noise floor high enough to hide the weak signal, although the weak signal would be fully visible if the strong signal were to disappear.

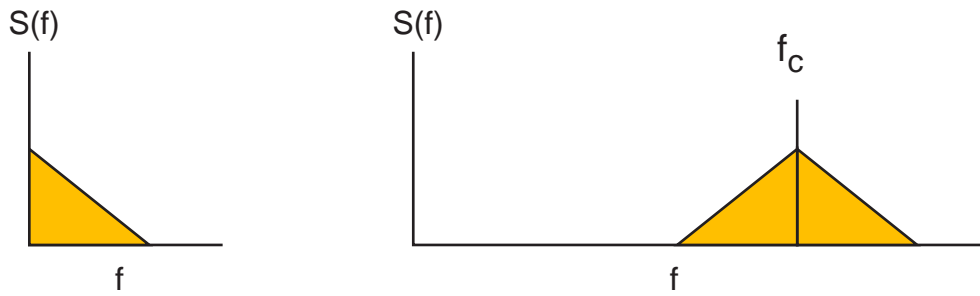


Figure 62. The up-conversion of baseband noise to “skirt” noise due to modulation at f_c .

This all sounds quite alarming, but in fact the fiber and solid state lasers that we propose to use for this program have REN noise that is confined to below 1 MHz. Thus, the upconverted noise puts noise skirts on the signal that can prevent the detection of signals that are really “close in”, but it will not affect signals located farther than 1 MHz away. Thus, for surveillance applications, laser excess noise plays no role. However, it does have ramifications for radar, where one often tries to get within 100 Hz of the carrier to detect Doppler shifts.

This problem of upconverted noise is not unique to AM. It shows up in IMDD, PM and FM systems also, although in the later two cases the origin is phase noise, not amplitude noise.

This brings us to another feature of equation (73): that there is no contribution from laser phase noise (other than the small amount in the noise \otimes noise terms). This is due to the fact that we are homodyne detecting in a way that maximizes the detection of amplitude variations, and minimizes the detection of phase variations. If we were to change ϕ to 0° , we would detect phase variations, and not amplitude variations. Hence, homodyne AM detection rejects phase noise of both the source and local oscillator lasers.

Interestingly enough, phase noise would probably not be an issue anyway, because we are planning on using the same laser for the source and LO. Because homodyne detection is the projection of one rotating vector onto another in an Argand diagram, equal phase noise causes the two vectors to wiggle together, so that the angle between the two is always the same, and the projection is unchanged. However, this means that the amplitude variations are also the same, so that the first two noise terms in equation (73) will add as amplitudes, not as powers, and the noise will be 3 dB worse than it would be for two independent lasers with slow phase locking.

Laser excess noise will degrade the SFDR when the spectral noise density of η multiplied by θ_o^2 exceeds the shot noise. For a 1 Hz bandwidth, this will be the case when the spectral noise is greater than -128 dBc/Hz. Examination of literature data for a few solid state lasers [1] shows that this will be the case for frequencies lower than 100-1000 Hz. For larger bandwidths, this threshold value shifts upward as the 1/3 power of bandwidth.

The lasers measured in [1] were used mostly for pulsed application, with no particular attempts being made to reduce amplitude noise. None, for example, employed REN noise suppression techniques. Hence, considerable improvement might be possible. However, it is important to emphasize, again, that this limitation exists only at frequencies very close to the

modulation frequency, and that at higher offset frequencies (> 1 MHz) equation (14) in section 2.2 holds unequivocally.

1. Scott, R.P., C. Langrock, and B.H. Kolner, *High-dynamic range laser amplitude and phase noise measurement techniques*. J. Selected Topics in Quantum Electronics, 2001. **7**(4): p. 641-655.

Appendix B: Effects of Finite Modulator Extinction

To analyze the effects of imperfect extinction, we shall represent the field amplitudes in each arm of the Mach Zehnder interferometer by ξ and $1-\xi$, with ξ being $1/2$ for perfect extinction. The field emerging from the modulator is the sum of the fields from each arm,

$$E(t) = E_1 + E_2 = E_s (\xi \cos(\omega_c t + \theta - \pi/2) + (1 - \xi) \cos(\omega_c t - \theta + \pi/2)) \quad (74)$$

When ξ is $1/2$, one gets the usual expression for fully suppressed carrier modulation,

$$E(t) = E_s \cos(\omega_c t) \sin \theta(t) \quad (75)$$

Applying the signal of equation (74) to the dual balanced detector described in section 2.2, and allowing the phase angle, ϕ , to have a slight offset, $\Delta\phi$, from its optimum value of $\pi/2$, gives

$$I_1 - I_2 = E_s E_{lo} \cos \Delta\phi \sin \theta(t) + (1 - 2\xi) E_s E_{lo} \sin \Delta\phi \cos \theta(t)$$

The first term is the normal signal, slightly reduced by the phase offset. The second term is the source of the second harmonic. Note that when $\Delta\phi$ is zero, one completely recovers the signal, regardless of any error in ξ .

Expanding $\cos\theta$ in a power series, one finds that the second harmonic signal is

$$I_{2\omega_m} = \frac{1}{4} (1 - 2\xi) 2\sqrt{I_s I_{lo}} \sin \Delta\phi \theta_o^2 \cos(2\omega_m t)$$

Repeating the procedure for the calculation of the SFDR outlined in section 2.2, one finds that the SFDR for a system having a small phase locking error, $\Delta\phi$ is given by

$$SFDR_{2\omega} = \frac{4}{|1 - 2\xi| \sin \Delta\phi} \left(\frac{I_s}{eB} \right)^{1/2} \quad (76)$$

The second harmonic begins to be a problem when this value is the same as that for IM_3 . This will happen when

$$|1 - 2\xi| \sin \Delta\phi = \left(\frac{eB}{I_s} \right)^{1/6} \quad (77)$$

Note that $|1 - 2\xi|$ is just the square root of the extinction ratio. Thus, for a 20 dB extinction ratio, $|1 - 2\xi|$ is 0.1.